Transactions

of the

ASME

MAY, 1949

VOL. 71, NO. 4

Transactions

of The American Society of Mechanical Engineers

Published on the tenth of every month, except March, June, September, and December

---

Published monthly by The American Society of Mechanical Engineers. Publication office at 20th and Northampton Streets, Easton, Pa. The editorial department is located at the headquarters of the Society, 29 West Thirty-Ninth Street, New York 18, N. Y. Cable address, "Dynamic," New York. Price $1.50 a copy, $12.00 a year for Transactions and the *Journal of Applied Mechanics;* to members and affiliates, $1.00 a copy, $6.00 a year. Changes of address must be received at Society headquarters three weeks before they are to be effective on the mailing list. Please send old as well as new address.... By-Law: The Society shall not be responsible for statements or opinions advanced in papers or...printed in its publications (B13, Par. 4)....Entered as second-class matter March 2, 1928, at the Post Office at Easton, Pa., under the Act of August 24, 1912....

Field Control of Oil-Field Pumping Equipment

By J. N. GREGORY,[1] LONG BEACH, CALIF.

The conditions faced by an oil producer in accomplishing continuous operation of surface pumping equipment without frequent attendance are outlined. The construction and application of devices to perform routine services on oil-well pumping equipment are described, and particular attention is directed toward eliminating the services which ordinarily must be performed at frequent intervals. The need for a preventive maintenance program is outlined and the instruments and tools which may be used in such a program are explained. Certain devices for use at problem wells are discussed, and the value of operating procedures which consider routine service, preventive maintenance, and corrective measures is indicated.

PRODUCING crude oil is unusual if it is compared with the usual production or assembly line of a manufacturing industry. The products are fluid and can be moved without manual handling, and there is no control of the quality of the product as it issues from the ground. On the other hand, the "production line" cannot be placed in an advantageous location nor arranged to make operation simple. Instead, oil is where you find it, and the oil producer encounters difficult terrain, adverse climate, widely spaced wells, and varied well conditions which pose many problems unique to the oil industry. It is the intent of this paper to describe some of the devices and programs which are effective in accomplishing continuous operation without breakdown or excessive servicing and maintenance costs.

Oil wells may be considered to fall into one of three categories: wells which should produce to capacity continuously; stripper wells which will produce to capacity without continuous operation; and wells with an allotment which is less than the productive capacity. Of these, the first are of the most concern since full-time operation at capacity is necessary if the greatest ultimate yield is to be realized, and to the cost of making repairs must be added the loss which may occur when production ceases. Thus the operator must select equipment capable of producing the well to capacity with enough allowance for normal wear so that an excess amount of time will not be required to replace worn parts, and he also must organize a service and maintenance routine which will assure continuous operation efficiently.

Continuous operation is not always necessary. In some fields the ability of the oil sands to give forth their contents is so limited that the capacity may be produced by intermittent operation, and careful planning will permit this to be done efficiently. If a field is produced at a lower rate than capacity, intermittent operation may be economical; however, the effect of a shutdown must be considered since many wells will be harmed if water enters the producing zone.

---

[1] Division Superintendent, Shell Oil Company, Inc.

Contributed by the Petroleum Division and presented at the Petroleum Mechanical Engineering Conference, Amarillo, Tex., October 3–6, 1948, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—PET-6.

If down time through failure is to be avoided, equipment must be kept in good operating condition, and imminent failure should be determined before a breakdown occurs. Three types of maintenance and repair service are required if the surface equipment is to be kept in first-class shape: (*a*) Routine service at frequent intervals to determine that the wells are pumping properly, to perform miscellaneous services such as adding oil and water to gas engines and lubricants to bearings, and to gage production rates; (*b*) preventive maintenance at infrequent intervals to determine engine performance and to inspect for worn or loose parts which may result in breakdown; (*c*) repair work as required to overhaul engines and replace faulty parts. There is no master plan which may be adopted to fit all conditions or fields, but if adequate preventive maintenance is inaugurated and devices are installed to do as much of the routine services as practical, it will be found that the time and work involved in pumping and down time from failure will be relatively small. Numerous devices have been developed to perform various services on oil-well pumping equipment, and they are listed with a brief description of their construction and function without regard to the novelty or desirability of exact application.

Special Devices for Oil-Well Pumping Equipment

Speed Regulation. If the maximum production is to be realized without excessive overpumping, close speed regulation is necessary. Electric-motor drives pose no problem other than proper design, since the induction motor is essentially constant speed. They do have the disadvantage that the speed cannot be readily adjusted to changing conditions. Gas engines equipped only with maximum-speed governors need frequent checking; the throttle must be adjusted manually to the desired speed and the setting may tend to change over a period of time. Also, until stable pumping conditions are reached, return trips to adjust the throttle often may be necessary.

A device known as a "Vacontrol," Fig. 1, has been developed to correct this. The Vacontrol is a throttle control actuated by the intake-manifold pressure. It consists of a cylinder piped to the intake manifold, and a piston connected through a linkage to the throttle. An adjustable spring controls the movement of the piston, and a lag valve, placed between the cylinder and the intake manifold, damps the fluctuations in pressure which tend to make the control surge. An idling screw and manual control lever complete the assembly.

It is the purpose of the control to adjust the throttle by the load instead of by the speed as is done with a conventional governor. In normal operation the spring is adjusted so that when the plunger is at the extreme end of its travel, the throttle for pulling the maximum load expected is obtained; thus any reduction in load will reduce the throttle. If the load is properly balanced, there will be no movement of the throttle; if not, the throttle will be adjusted to load changes as they occur. If extreme unbalance exists, the lag valve prevents the piston from moving rapidly enough to keep up with the changing load and the engine will die.

The Vacontrol serves four purposes:

1 It smooths out erratic motion.

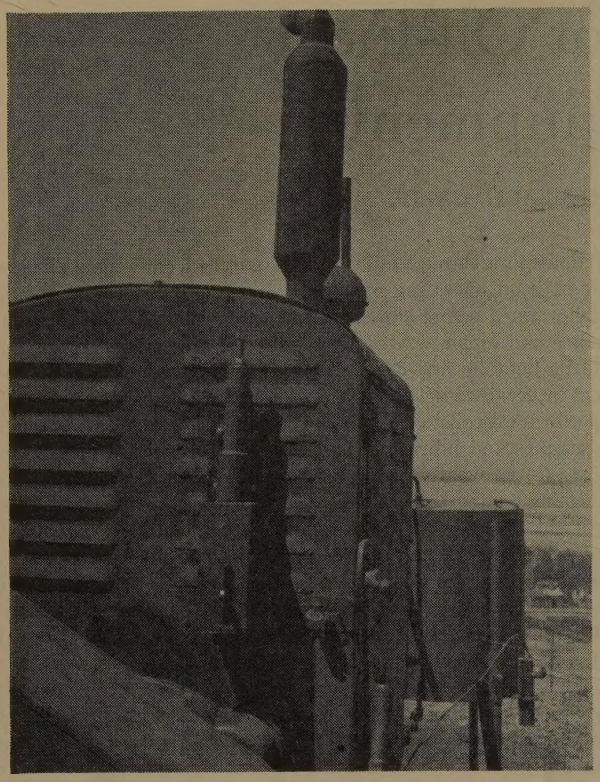


FIG. 1 THROTTLE CONTROL ACTUATED BY INTAKE-MANIFOLD PRESSURE

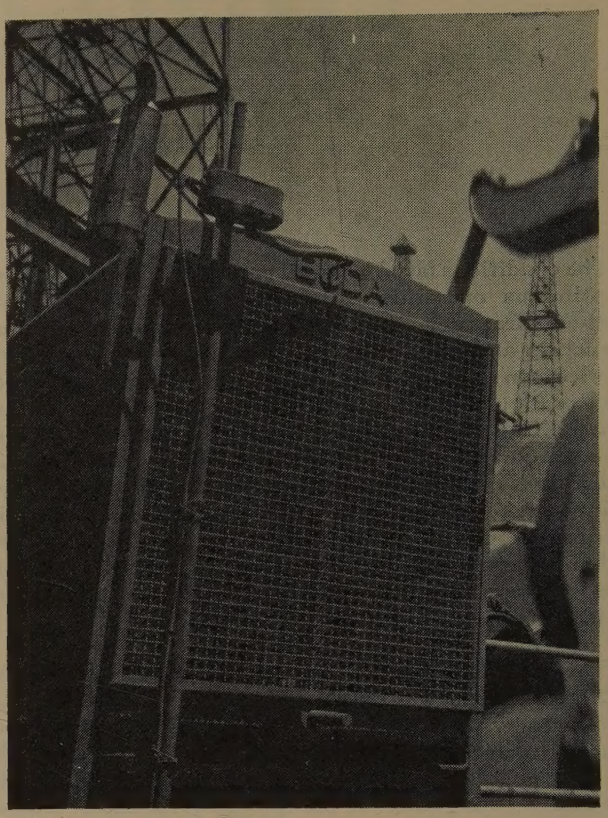


FIG. 2 RADIATOR WATER-LEVEL FLOAT CONTROL

2 It serves as a safety shutoff when the load becomes unduly out of balance.

3 It permits the engine to be declutched and re-engaged without manually adjusting the throttle to the proper operating speed.

4 If a well has been shut down and must be pumped up (such as occurs after a pulling job), the Vacontrol will adjust the throttle to the load automatically until stable conditions are reached. This avoids return trips to the well to adjust the engine speed.

Lubricating and Cooling Systems. The regularly equipped gas engine requires frequent additions of oil and water. When the work is performed manually, there is a tendency to overfill the crankcase, which is wasteful in itself, and although the time involved in servicing one engine may be small, collectively a number of engines consumes a surprising amount of time. A practice which is becoming widespread is the storage of adequate amounts at each engine so oil and water can be added automatically, maintaining constant levels. There are suitable float valves available for this purpose; one of them suitable for the control of water in the radiator is a humidifier control valve which is snap-acting and self-cleaning, Fig. 2. Working parts of the valve are above the water level to reduce corrosion and scale accumulation. An oil-float valve has been developed specifically for maintaining an accurate oil level; it is equipped with a screen filter, visible sludge chamber, sight glass, and vacuum breaker for engines which operate with a vacuum in the crankcase, Fig. 3.

Another type is a vacuum-controlled oil-level regulator, Fig. 4. It consists of an airtight drum, piped from the bottom to the engine crankcase. An air-vent tube connected to the top is adjusted so its free end touches the oil at its proper level; when the level drops below the vent-tube opening, air is admitted to the top of the drum, allowing oil to drain into the crankcase until the vent-tube opening is closed by the oil surface. Vacuum from

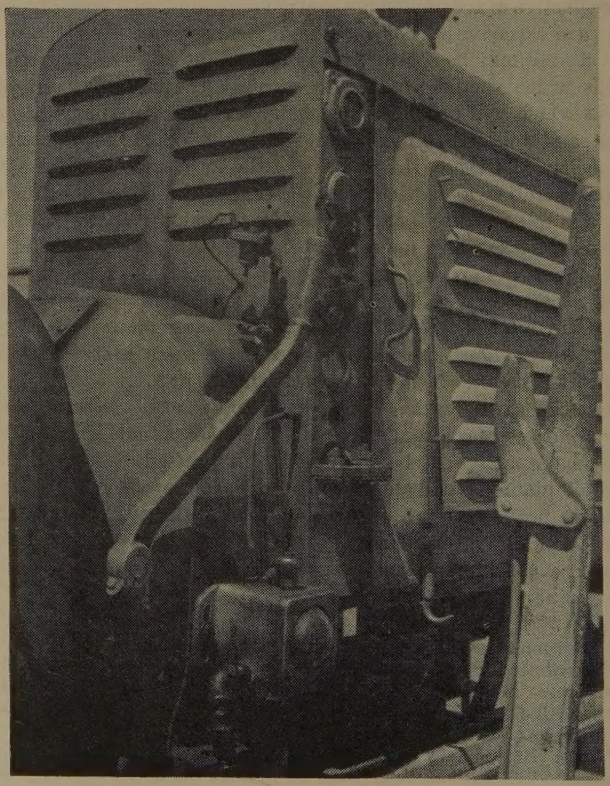


FIG. 3 CRANKCASE OIL-LEVEL FLOAT CONTROL

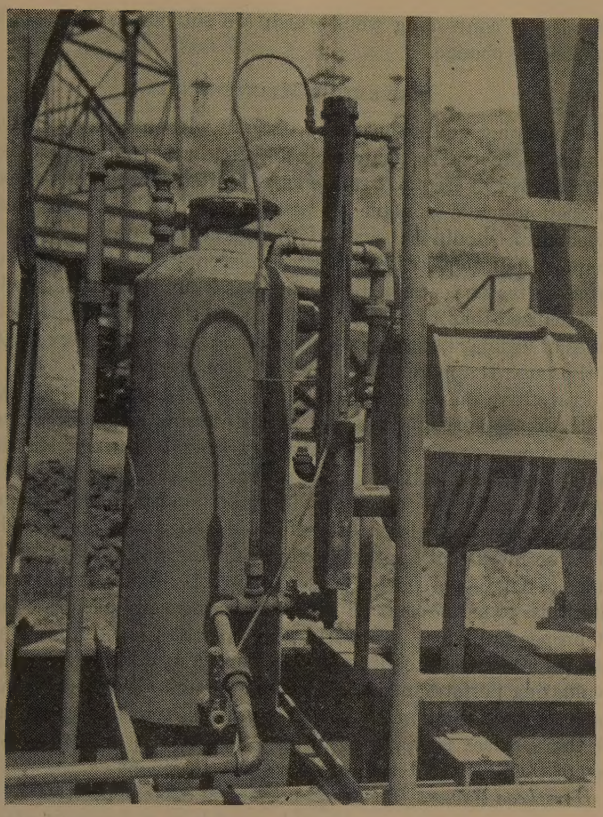


FIG. 4 CRANKCASE VACUUM OIL-LEVEL CONTROL

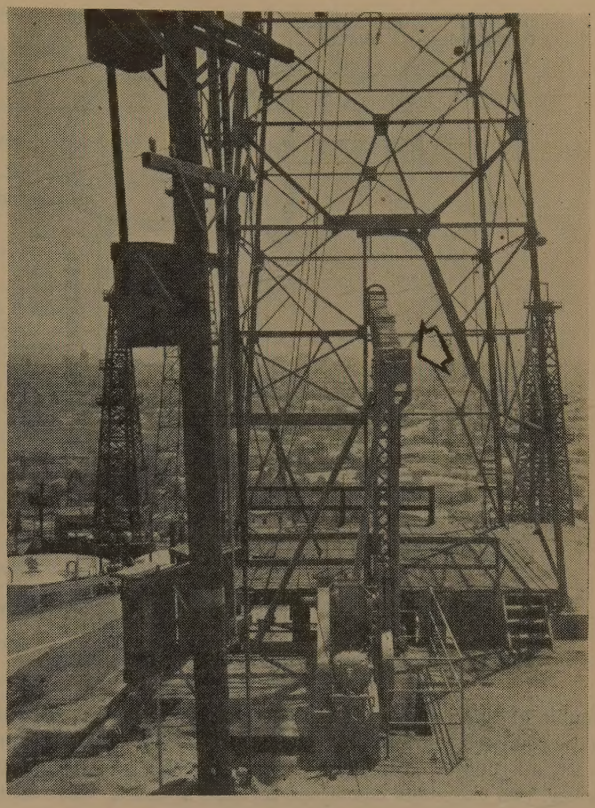


FIG. 5 APPLICATION OF GRAPHITE-IMPREGNATED PLASTIC OILLESS BEARING

the intake manifold connected to the top of the drum, with a restirctor, draws a very small quantity of air continually through the air-vent tube; this compensates for atmospheric-temperature changes and draws any excess of oil which might find its way into the crankcase back into the drum.

Such devices, and the storage of an adequate quantity of lubricating oil at the well, eliminate the time-consuming addition of oil and water at frequent intervals, and tend to reduce oil consumption and contamination.

Bearings. Properly designed bearings in the pumping unit may go without service for long periods of time if an adequate oil reservoir is built into the bearing assembly and oil seals are maintained. A bearing requiring no lubrication has been used in some areas with success, Fig. 5. It consists of a graphite-impregnated plastic block with a long-fiber-asbestos binder, molded under 6000 psi. Its compressive strength is some 11,500 psi determined on a 3-in. unsupported cube, and it has a Brinell hardness of about 84. Some of these have been in use for 8 years without failure or measurable wear. Their particular use in oil-field equipment has been confined to the overhead members such as the upper Pitman and Sampson-post bearings which are difficult to lubricate.

Automatic Shutoffs. If a pumping unit is to operate without frequent service and inspection, safety devices to shut the unit down in the event of failure are as necessary as devices to eliminate frequent lubrication. Electric motors may be protected with overload heater relays, although a more foolproof and sensitive device is the "Thermoguard" which consists of a thermostatically operated switch built into the motor winding. It has the advantage of breaking the circuit if the motor itself overheats from any reason such as fire or "single phasing."

Gas engines may be protected by oil-pressure and thermostatically controlled switches which shut the engine down in the event necessary engine auxiliaries fail. Other shutoffs, such as inertia switches on walking beams, and pressure-operated switches of air counterbalance units, are also available.

Stripper wells which operate part time may be shut down by clockwork if they may be operated on a fixed schedule. Another device which has found some application consists of a bucket on a switch lever hung under the outlet of the flow line. A small hole in the bottom permits oil to flow out, but at a slower rate than it enters if the well is pumping. When the well pumps off, the bucket empties and opens the switch controlling the prime mover.

Automatic Starters. Stripper wells which operate part time may be started by clockwork if they are pumped electrically. In general, a group of wells operated under this system should be operated on a staggered schedule in order to keep the power demand low.

Power failure, even for short duration, will shut down an electrically operated lease. If the line starters are connected to start the motors when the current comes back on, the high starting currents may be too severe for the power system. To prevent this, delay switches must be used. Several types are available. Among the more common are motor-driven cams, dashpots, and thermostats. In each, the principle is to delay the application of current to the holding coil for a predetermined time after the circuits have been energized. No known attempts to start gas engines automatically have been made It is doubtful that automatic engine starting will ever find much application, since shutdowns usually indicate trouble in the unit itself.

Counterbalancing. The need for proper counterbalancing is accentuated if close attention is paid to economical equipment selection and operation. The need for readily variable counterbalancing does not appear to be universal, but certain wells cause

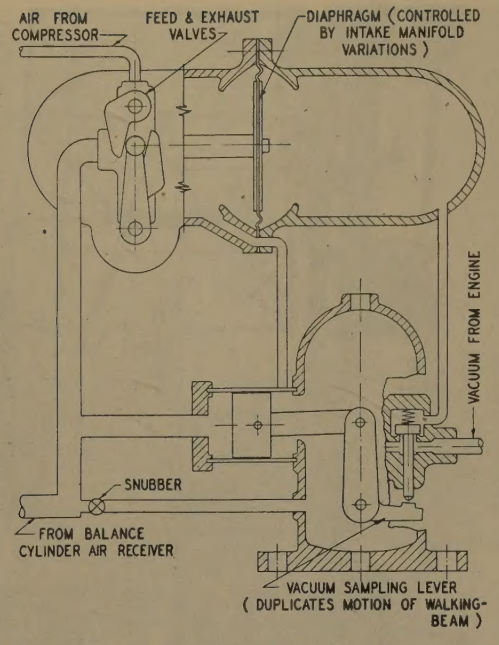


FIG. 6 DIAGRAM OF AUTOMATIC COUNTERBALANCE

trouble of this nature. It is interesting to note that an automatic counterbalance has been developed which is easily adapted to any air-counterbalanced unit, Fig. 6. With a gas-engine prime mover, two chambers connected through a timing valve to the engine manifold are used. The timing valve is actuated so one chamber is exposed to the manifold pressure during part of the upstroke, the other during the counterpart of the downstroke. If perfect counterbalance exists, the pressures in the two chambers are equal; if not, a diaphragm between the two chambers actuates a valve in the air-counterbalance system, varying the pressure until proper balance exists. If an electric-motor drive is used, solenoids are substituted for the vacuum chambers. One of these devices has demonstrated its ability to maintain proper balance even when the polish rod parted.

TROUBLE CALLS

It would be ideal if it were never necessary to visit a well which was performing properly. Achievement of such a goal in the near future does not appear likely, but an approach lies in radio transmission, which has been tried in certain areas. Transmission of all performance data would be quite complex, so efforts have been confined to the transmission of poor performance. An arbitrary selection of standards is made, and transmission occurs only when the standards are not maintained. The signal transmitted indicates only that failure has occurred and that a visit to the well is necessary.

The equipment has three components; an initiator such as a pressure gage with electrical contacts set at the upper and lower limits of satisfactory values; a transmitter having a crystal-controlled oscillator and a tone modulator; and a crystal-controlled superheterodyne receiver, equipped with decoding and indicating devices. The keying mechanism in the transmitter is automatic in its cycle, which requires a maximum of 5 sec, and 400 combinations of pulses are possible. The short transmission time and remote likelihood that two wells will fail simultaneously allow as many as 400 wells to be controlled. The receiver, with its filters, relays, and decoding devices, may signal with lights, buzzers, or graphically.

Such radio transmission has been used to indicate pressures at flowing wells and gas-engine performance successfully, and offers promise in controlling wells which are subject to failure when conditions change.

PREVENTIVE MAINTENANCE

If pumping equipment is to operate without failure, a well-planned preventive maintenance program must be devised. Particularly, if frequent servicing is limited to the very cursory inspection made possible by the devices outlined, the entire mechanism should be serviced thoroughly and inspected at infrequent but regular intervals. Experience indicates that such intervals may be extended to as much as 2 months, although when it is realized that a period of 1400 hr is equivalent to some 40,000 miles of ordinary automobile operation, the need for service that is thorough is apparent.

Preventive maintenance should include inspection of all parts of the machinery; lubrication of all wearing surfaces; adjustment, repair, or renewal of faulty parts such as bearings, clutches, belts, horsehead cables, and ignition, cooling, and lubricating systems; testing of engine conditions to determine performance; and a report of all services performed and conditions noted to allow an intelligent appraisal of the need for major repairs. The value of a written report cannot be overemphasized; it serves as a check list to assure that all necessary services are performed and at proper intervals; it reports poor operating conditions which call for a change in equipment, and provides a history useful in evaluating performance and need for repairs. A simple service memorandum is sufficient, and the desirability of recording most of the services is apparent, Fig. 7. However, discussion of a few of them with the tools used may be pertinent.

Horsehead Cables. The usual horsehead cable, babbitted to the carrier bar, may give considerable trouble in a corrosive atmosphere if it does not have proper care. Periodic lubrication will lengthen the life, and a galvanized cable is now available which promises to resist corrosion. At best, the wire line adjacent to the babbitted socket is subject to early failure due to the effect of heat while attaching the socket. One carrier bar now available avoids this; it is so formed that the wire line may be carried around it in back of the polish rod while maintaining a vertical pull to avoid bending in the polish rod, Fig. 8. Thus the horsehead cable may be made continuous by joining the two ends with wire line clamps. Cables so installed may be replaced easily, and by moving them from time to time, bending fatigue in one spot may be reduced and the life increased.

Fuel Adjustment. The carburetor adjustment is quite important. On natural gas an engine with a lean mixture sounds smooth and sweet, but it will be found that the excess oxygen and high cylinder temperatures will burn valves and oil and clog the ring slots. Temperatures may be reduced appreciably, avoiding this trouble, if the carburetor is adjusted to an air-fuel ratio of about 12.8 to 1. An exhaust-gas analyzer is the most practical method of determining this; one of the types which is rugged and easy to use is an instrument which applies the thermal conductivity principle. In this, two spirals of platinum wire are enclosed in separate cells encased in a solid metal block to insure equal temperatures. One cell is filled with a standard gas with known characteristics; the exhaust gases are passed through the other cell. Each coil forms one arm of a Wheatstone bridge through which a constant current flows, heating the spirals. The loss of heat through the gases surrounding the spirals is dependent on the thermal conductivity of the gases. Hence by measuring the resistance, the temperature difference may be determined which is a measure of the difference in thermal conductivities. The air-fuel ratio is proportional to the thermal conductivity, so the Wheatstone-bridge galvanometer reading is a direct indication of the mixture.

Engine Performance. The speed of the engine is measured to

Unit No._____ PUMPING UNIT AND GAS ENGINE SERVICE MEMORANDUM Date_____

Location_____ By _____

Work Symbols: 00 = Examine: 0 = Repair: X = Adjust + = Renew: — = Clean

Head	Gear Box	Compression:	Carburetor
Cylinders	High Speed Shaft	#1 #2 #3	Regulator
Pistons	Low Speed Shaft	#4 #5 #6	Air Cleaner
Pins	Brake	Oil Pressure	Governor
Rings	Idler Pulley	Oil Switch	RPM
Rods	Crosshead Bearing	Oil Temperature	Vacuum-load
Main Bearings	Outboard Bearing	Oil Filter	Vacuum-Idle
Crank Shaft	Drive Belts	Oil Screen	Radiator
Cam Shaft	Wrist Pin	Oil Change	Water Temp.
Oil Pump	Pitman Bearings	Oil Consumption	Water Pump
Timing Gears	Sampson Post Bearing	Magneto	Fan
Valves	Walking Beam	Spark Plugs	Fan Belt
Muffler	Horsehead Cable	Wiring	Hose
Clutch	Base & Fndtn. Bolts	Starter	Lube Motor
	Housekeeping	Generator	Lube Pumping Unit

FIG. 7 PREVENTIVE MAINTENANCE CHECK LIST

determine if it lies in the optimum range, and from vacuum readings the horsepower may be determined. The finest maintenance program will be of little avail if engines are not properly loaded, and improper application may be the key to many engine failures.

FIG. 8 HORSEHEAD CABLE SECURED WITH WIRE LINE CLAMPS

The water temperature, oil-temperature, oil-consumption, and compression values are excellent guides to engine wear. Another tool with considerable value is a periodic analysis of the crankcase oil. Samples are withdrawn from the bottom of the crankcase and analyzed for such properties as viscosity, sludge, water content, and foreign matter from which may be determined high piston temperatures, clogged oil-ring slots, dirty air cleaners and oil filters, contaminated oil-can measures, improper ignition, bearing wear, and high and low crankcase-operating temperatures, Fig. 9. Such a service provides an excellent control but is of greater value to the mechanic who uses it intelligently as a tool. One value of such analyses is indicated by the experience that many engines operating under optimum condi-

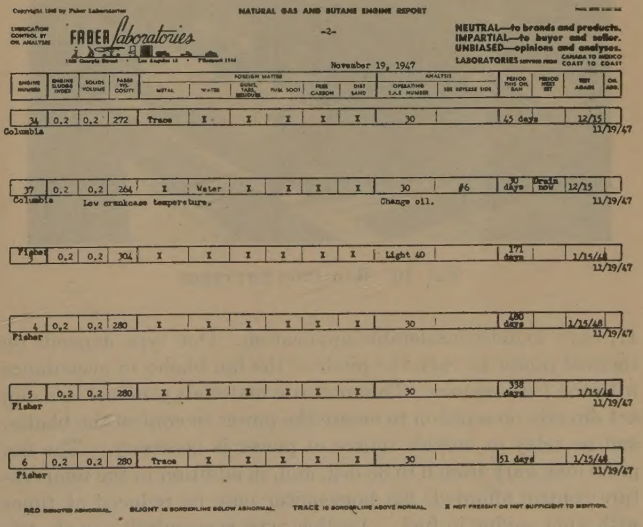


FIG. 9 CRANKCASE-OIL-ANALYSIS CHART

tions have not required repair or an oil change in more than 2 years when dependence has been placed on service records and oil analyses. However, of more value is the number of engines which have been discovered with serious faults which undoubtedly would have broken down had dependence been placed only upon a cursory inspection and service. Oil analyses also may be applied to gear cases to determine oil contamination and wear of bearing surfaces.

Problem Wells

In spite of proper application of gas engines, problem wells and conditions invariably will be encountered, one of the most serious of which is temperature variation in the crankcase, leading to condensation and contamination of the lubricant. This condition is often found where equipment is subject to climatic variation, and several devices have been developed to assure a more uniform temperature.

Radiator Shutters. Radiator shutters, Fig. 10, control the supply of air to the radiator and are usually operated by a vacuum cylinder through a thermostatic relay mounted between the engine head and radiator. They are quite effective in maintaining a uniform temperature around an engine which is properly housed.

Thermally Controlled Fans. Thermally controlled fans, Fig. 11, have found considerable application. One type depends on thermal power to vary the pitch of the fan blades in accordance with the temperature. Thermostatic materials confined in a cup act directly on a piston to create the power to control the blades, and no relay or outside source of power is necessary. The fan pitch may vary from 0 to 35 deg, and, in addition to the temperature control afforded, fan horsepower may be reduced at times with some saving in fuel. Another type is essentially an electric clutch. The rotating speed of the fan varies according to the electric current in the field which is thermostatically controlled through a relay. A source of direct current is necessary for this device.

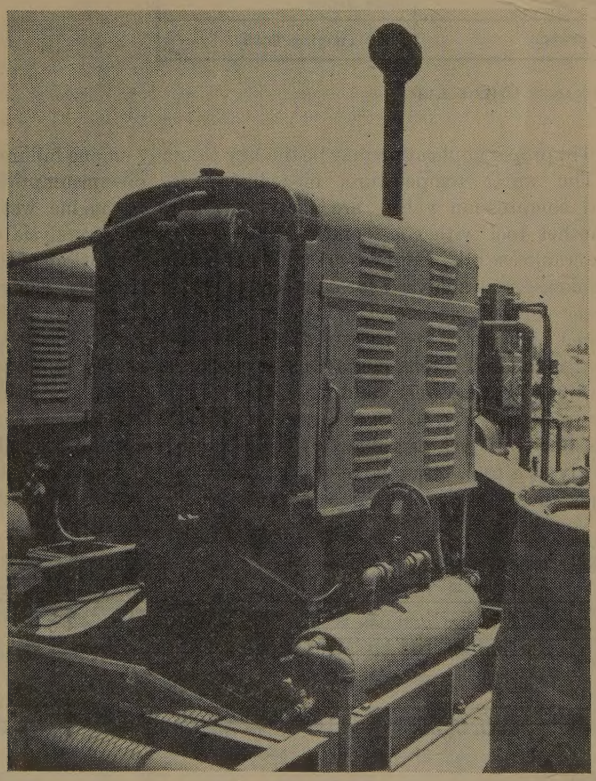


Fig. 10 Radiator Shutters

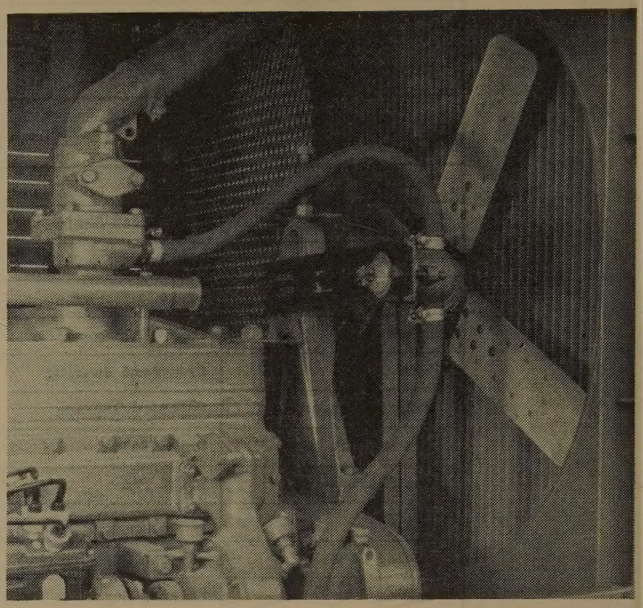


Fig. 11 Thermally Controlled Variable-Pitch Fan

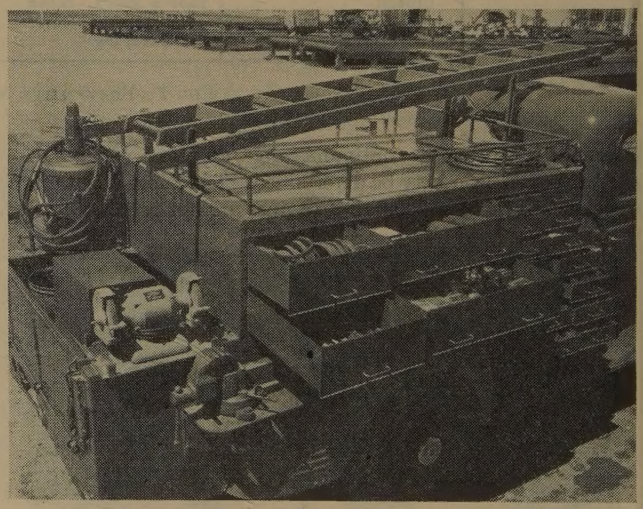


Fig. 12 Fully Equipped Tool Truck

Vapor-Phase Cooling. Vapor-phase cooling was designed to eliminate hot and cold spots around the cylinder walls. The jacket water is maintained at or near 212 F, and heat is withdrawn from the engine by flashing the water into steam at its exit from the engine, condensing it to water in the radiator, and recirculating it at a temperature near boiling.

Jacketed Crankcase. In order to obtain more uniform temperatures, crankcases occasionally have been jacketed to be included in the cooling-water system. It is understood that this has been found particularly useful in areas where contamination of lubricating oil from the fuel is experienced.

Preventive and Corrective Maintenance

Since wide spacing and remote locations are conditions usually faced by the oil producer, traveling time and transportation of materials are major factors in any service program. Furthermore, it is recognized that any successful maintenance program

must be directed toward both preventive and corrective measures. It is not desirable to spend much time on equipment that is operating properly, so the preventive phases should be made as routine and simple as possible. Provision of tool trucks, Fig. 12, or even small traveling shops especially designed for the purpose will simplify the routine portion of the work, and adequate and properly arranged tools and repair parts on the job may avoid long trips to find parts.

Haphazard installations of special devices on all pumping units would not be sound, nor would the inauguration of a preventive maintenance program without both consideration of the conditions involved and the benefits to be gained. The cost of inspection and preventive maintenance must be balanced with the cost of down time and repairs avoided, and each area must be evaluated properly. It has been demonstrated that even in fields where complex conditions exist and multiple sizes and types of units are used, pumping service of 40 to 60 wells per man is within reason. Also, a preventive maintenance program under the same conditions can be carried out at the rate of 50 units per man with 20 per cent of his time allowed for corrective measures and emergency repairs. It is emphasized that any program should be directed toward corrective measures; repairing a failure is usually more expensive than correcting a fault. Preventive maintenance can be made routine whereas failures are unpredictable, and, if primary work is routine, more attention can be directed to the corrective measures which in the end will assure the producer of steady, economical operation without costly failure.

Operating Characteristics and Requirements of Drawworks Brake Equipment

By R. G. De La MATER,[1] PARKERSBURG, W.VA.

In this paper an analysis is made of the operating characteristics, capacity, and limitations of the mechanical friction brakes and of the hydraulic and electric types of supplemental dynamic brakes used on drawworks. Consideration is then given to the relative advantages and disadvantages of various methods of drawworks operation while running in the drill pipe. The speed-load characteristics of the drawworks, for a preferred method of operation, are then compared with the speed-capacity characteristics of the several available brake systems. Consideration is also given to the requirements of the driving means for supplemental dynamic brakes.

Introduction

UP to the time that the majority of oil wells did not exceed 3000 to 4000 ft in depth, the design and operation of the brake equipment of drilling hoists did not present any serious problems. However, with the drilling of wells to ever-increasing depths beyond 4000 ft, it was found that, due to their inherent limitations, mechanical friction brakes were becoming increasingly inadequate for either safe or economical operation.

In 1932 the hydrodynamic brake was introduced as an adjunct to the mechanical friction brakes of drawworks and rapidly became recognized not only as a solution to the braking problem when drilling to depths of 5000 ft or more but also as a distinct advantage when drilling to lesser depths. The brake system comprising the combination of mechanical brakes with some type of supplemental dynamic type of "energy dissipator" has now been adopted practically universally for all but the lightest of drawworks.

Because of the great depth to which wells are now being drilled, as well as the present great interest in speedier, more efficient, and more economical drilling operations at all depths, it is essential that the characteristics and limitations of the brakes now available for drawworks be examined in the light of the service to which they are being subjected. Such is the purpose of this paper.

Mechanical-Friction-Brake Limitations

The manufacturer properly provides on each size drawworks the largest and most efficient mechanical-friction-brake assembly that the design limitations and operating conditions of the drawworks will permit. However, the capacity of these brakes to dissipate energy, in the form of heat, has definite limitations which for intelligent operation should be understood thoroughly.

Friction-brake lining is made up of asbestos, a filler, and a binder. If asbestos, a fibrous crystalline form of magnesium silicate, is heated to 500 F for sustained periods, or subjected to higher temperatures for shorter periods, its water of crystallization distills off, thus reducing it to a white dust. Temperatures of 500 F also will cause oxidation and deterioration of organic binders, and generally adversely affect the frictional characteristics of the lining.

With mechanical friction brakes the energy to be dissipated is converted into heat at the friction surface. As illustrated in Fig. 1, a temperature differential, the equivalent of voltage drop in the case of electrical circuits, is required to conduct this heat through the rim away from the friction surface, through any scale deposit or other fouling on the inner surface and through a

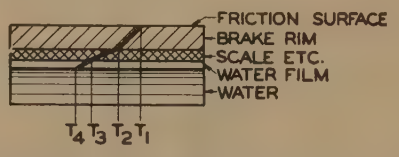


Fig. 1 Temperature Gradient Through Rim of Water-Cooled Mechanical Friction Brakes for Drawworks

water film to the circulated cooling water. As a typical example, if T_1, the temperature at the friction surface, is 500 F, and if T_4, the temperature of the cooling water, is 80 F, T_2 at the inside surface of the rim may be 372 F, and T_3 at the inside surface of the scale deposit may be 245 F. The temperature differential through the rim causes the friction surface to be in compression and the inner surface to be in tension. The resulting stresses at the rim surfaces cannot be determined by exact analysis but approximate calculations indicate, for a temperature differential of 128 F, that they may be in the order of 30,000 psi.

It is obvious, therefore, that the physical limitations of both lining and rims dictate brake operating temperatures not exceeding 500 F, and many authorities recommend temperatures of 400 F or less.

Calculations for brake capacities are frequently based upon arbitrary values of allowable unit pressure between lining and rim or horsepower that may be dissipated per unit of friction surface. These are not satisfactory as a basis for determining the capacity of drawworks brakes, unless there is also taken into account the all-important consideration of providing for the dissipation of the heat. A proper approach is to consider the problem on the basis of thermal capacity to dissipate heat, and while the scope of this paper does not permit the development of such an analysis, the results are shown in Figs. 2 and 3.

Modern drawworks recommended for 6000-ft drilling are equipped with two water-cooled friction brakes which average 43 in. diam × 8 in. face and 15 sq ft total circumferential area. Specifications generally list a supplemental energy-absorption-type brake as optional, but specifically recommend its use for depths of about 5000 ft. The diagram in Fig. 2 clearly indicates that for this size rig and to hold mechanical-friction-brake operating temperatures between 400 and 500 F, 4250 to 4500 ft is the maximum depth to which pipe should be run in without the use of a supplemental energy dissipator.

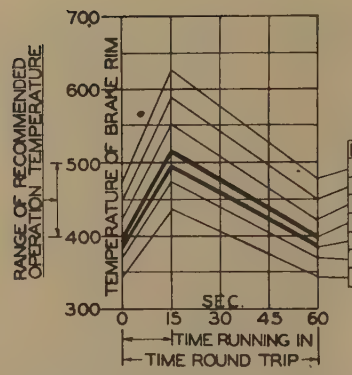
It is important to point out that during the 15 sec of actual running-in time for a stand of pipe, and during which time all the energy is converted into heat by the brakes, only about 20 per cent of the heat is actually conducted to and carried away by the cooling water. The balance, approximately 80 per cent of the

---

[1] The Parkersburg Rig & Reel Company

Contributed by the Petroleum Division and presented at the Petroleum Mechanical Engineering Conference, Amarillo, Tex., October 3–6, 1948, of The American Society of Mechanical Engineers.

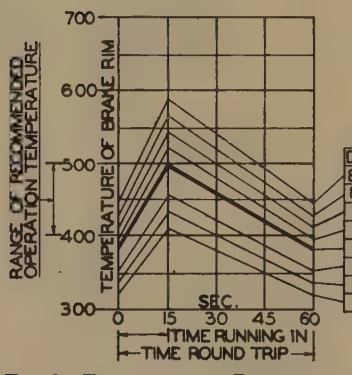
Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—PET-3.



DEPTH	BRAKE RIM TEMP. AVER.	MAX.	MIN.	WT. * SUSP. LOAD	WT. SUSP. LOAD / BRAKE RIM AREA (15 SQ. FT.)
6000	551	627	475	120380	8030
5500	519	589	449	111380	7430
5000	486	551	421	102380	6830
4500	454	513	395	93380	6230
4250	438	494	382	88880	5930
4000	422	475	369	84380	5630
3500	390	437	343	75380	5030

* AIR WEIGHT

FIG. 2 TEMPERATURES DEVELOPED AT FRICTION SURFACE OF MECHANICAL-FRICTION-BRAKE RIMS WHEN RUNNING-IN DRILL PIPE WITH MECHANICAL FRICTION BRAKES ONLY
(For drawworks recommended for 6000-ft drilling.)



DEPTH	BRAKE RIM TEMP. AVER.	MAX.	MIN.	WT. * SUSP. LOAD	WT. SUSP. LOAD / BRAKE RIM AREA (25 SQ. FT.)
8500	518	590	445	184760	7390
8000	498	568	429	175760	7030
7500	479	545	413	166760	6670
7000	460	522	398	157760	6310
6500	440	499	382	148760	5950
6000	405	457	353	132380	5295
5500	386	434	337	123380	4935
5000	366	411	321	114380	4575

* AIR WEIGHT

FIG. 3 TEMPERATURES DEVELOPED AT FRICTION SURFACE OF MECHANICAL-FRICTION-BRAKE RIMS WHEN RUNNING-IN DRILL PIPE WITH MECHANICAL FRICTION BRAKES ONLY
(For drawworks recommended for 12,000-ft drilling.)

heat, is stored in the rim, with the result that there is an appreciable increase in rim temperature. The remaining 45 sec of the round trip are required for this stored heat to be conducted to and carried away by the cooling water with an accompanying decrease in rim temperature. As shown in Fig. 2, for the 4250-ft stand and a surface rim temperature at the start of 382 F, the temperature will increase to 494 F by the time the stand is run-in and then drop back to about 382 F before the next stand is run. The high rim temperature itself, averaging 438 F at 4250 ft, is required to provide the necessary temperature differential for conducting the heat from the friction surface to the cooling water. As the weight of suspended load becomes greater and the energy to be dissipated increases, the temperature must also increase.

Drawworks now recommended for 12,000-ft drilling are equipped with two mechanical friction brakes averaging 54.7 in. diam × 10.5 in. face and 25 sq ft total circumferential area. In addition to the mechanical brakes, these drawworks are equipped universally with a supplemental dynamic energy dissipator. The diagram in Fig. 3 indicates that with this size rig, and under normal conditions of operation, 6500 ft is the maximum depth to which pipe should be run-in without the use of the supplemental energy dissipator.

The tabulations accompanying Figs. 2 and 3 show the weight of suspended load used in the calculations from which the rim temperatures were determined and the ratio between these weights and the brake-rim area. On the basis of these ratios the following thermal rating for mechanical friction brakes, in terms of weight of suspended hook load per square foot of circumferential brake rim area, is indicated:

For average conditions of operation, assuming 15 sec for lowering a 90-ft stand and 1 min per round trip, and a buoyancy factor for the mud of 0.85, the maximum ratio of hook load (air weight) to drawworks brake-rim area is 5900 psf.

Corresponding calculations for other speeds of operation demonstrate a very important factor, i.e., that without exceeding safe operating temperatures, for each increase in time of running in a 90-ft stand and round trip by 3 sec, the allowable hook load is increased by approximately 300 lb per sq ft of brake surface with the allowable depth of operation increased accordingly. For faster speeds of operation the allowable hook load is decreased accordingly.

DYNAMIC ENERGY DISSIPATORS

In contrast with the capacity limitations of mechanical friction brakes, dynamic-energy-dissipating devices are now available in a range of sizes and capacities to meet every drilling condition and requirement of the present or foreseeable future.

There are two basic types of dynamic energy dissipators, the "hydrodynamic" and the "electrodynamic" (eddy-current). Both types may be driven from the drum shaft in essentially the same manner, and both develop resistance only when in motion. Beyond these two points their characteristics are generally quite different.

Neither of the two types can actually stop and hold a load, functions implied at least by the term "brake," and consequently neither type can entirely replace the regular mechanical friction brakes. Hence the author has chosen generally to designate both types as "supplemental energy dissipators," and to bring

out in later discussion the difference in degree to which they actually are supplemental.

With the hydrodynamic type, resistance is created exclusively by fluid friction and agitation of water circulated between the veined pockets of the rotor and stator elements. The conversion from mechanical energy to heat takes place directly within the water itself. The water required is generally supplied by the regular service pump for the rig. During the running-in of a stand, and while revolving in the direction to create resistance, this type induces its own continuous inflow of fresh, cool water to displace the heated fluid. As a consequence the energy is dissipated at the same rate it is generated by the descending load. It is not necessary for the brake to be rotated while raising the empty blocks. Nor in any case is water circulation required during that period.

The resistance at any speed is adjustable by regulation of the volume of water being circulated through the veined pockets of the rotor and stator. Depending upon the type of installation, infinitely close regulation or regulation in any desired number of steps may be obtained by simple valve arrangements to vary the rate of flow or vary the head of water in a supply standpipe.

For any setting of the resistance control the horsepower capacity increases approximately in proportion to the cube of the speed; for example, if the speed is doubled, the horsepower resistance is increased 8 times. In starting from rest or while in operation, the resistance is immediately responsive to any change of speed or setting of the resistance control. However, for any sudden appreciable change in either, there may be a time lag of as much as 1 or 2 sec during which the water circulation and attending resistance rapidly but smoothly re-establishes itself.

If for any reason the water supply should be cut off during the running-in period, the water in the casing, and until completely evacuated, will keep the speed sufficiently in check to give ample time for the mechanical brakes to be applied.

With the electrodynamic type, resistance is created by eddy currents produced by rotating an iron armature drum through the magnetic field of stationary electromagnets. The heat thus generated at the surface of the drum is dissipated by convection to the cooling water circulated around the coils and drum. To effect this heat transfer to the water there must be a time element and a temperature differential between rim and water, the magnitude of each depending upon the amount of energy being dissipated and the rate of flow of cooling water. To dissipate the heat approximately as generated, when operating at maximum capacity, would necessitate circulating water around the rotor at a rate of about 200 gpm. However, since the flywheel effect of the rotor is not appreciable, the rotor is turned in the opposite direction during the raising of the empty blocks and by circulating water continuously at only about 25 per cent of the maximum rate otherwise required, the heat stored in the rotor during the running-in of a stand is dissipated during the time of the round trip without exceeding allowable operating temperatures. A signal horn is provided to blow in the event the water circulation is not sufficient to keep the drum temperature within safe operating limits.

The resistance at any speed is adjustable by regulation of the direct current used to excite the field coils. For maximum resistance only about 70-amp or 7-kw excitation is required, and this is usually available from the stand-by light plant of the rig. In the case of an alternating-current supply, a rectifier is used to convert the current to direct current. The manufacturer has considerable latitude of choice in selecting means of regulating the field-coil current. The control system now used for either a-c or d-c supply is a four-contactor drum switch, giving a choice of four degrees of resistance, and as it is an easy matter to move through the various contacts to control the load at the desired speed.

For any setting of the resistance control the horsepower capacity increases approximately in proportion to the 1.25 power of the speed; for example, if the speed is doubled, the horsepower resistance is increased $2^1/_2$ times. In starting from rest or while in operation, the resistance is immediately responsive to any change in speed, and for any sudden appreciable change in the setting of the resistance control, there is but a fraction of a second lag in the excitation of the field coils and corresponding change in resistance.

Fig. 4 illustrates the speed-horsepower characteristics of two approximately comparable sizes of electrodynamic and hydrodynamic energy dissipators when operated with the maximum setting of their respective resistance controls. These curves are also characteristic for each type with lower settings of their resistance controls.

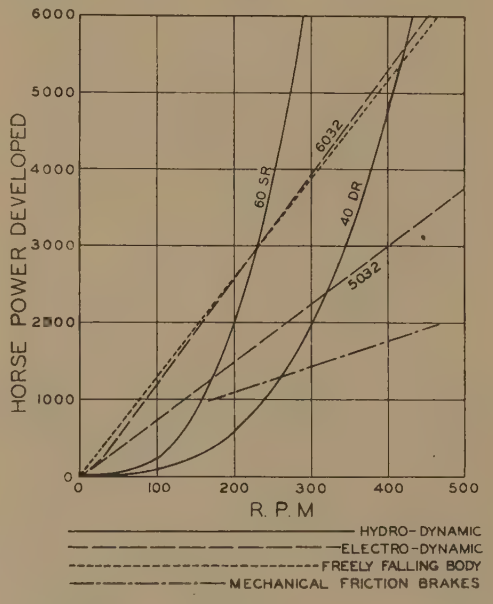


FIG. 4 COMPARATIVE CHARACTERISTIC CURVES OF ELECTRODYNAMIC AND HYDRODYNAMIC SUPPLEMENTAL ENERGY DISSIPATORS AND OF MECHANICAL FRICTION BRAKES

It will be noted the capacity of the 40 DR hydrodynamic equals that of the 5032 electrodynamic at 320 rpm and the 6032 at 420 rpm; also, that due to the difference in speed-horsepower characteristics, its capacity increases more rapidly than the 5032 above 320 rpm and decreases more rapidly below 320 rpm. The capacity of the 60 SR hydrodynamic equals that of the 5032 electrodynamic at 170 rpm, and of the 6032 at 230 rpm. In the high range of speeds, both the electrodynamic and hydrodynamic types have capacities well in excess of the requirements of any drawworks application. At the very low speeds of operation the capacity of the hydrodynamic type is negligible.

Further to illustrate the capacity characteristics of the two types, there is also shown in Fig. 4 the speed-horsepower curve for a freely falling body of sufficient weight to develop the same horsepower at 230 rpm, as the hydrodynamic and electrodynamic energy dissipators. The curve of the electrodynamic type practically parallels this curve whereas the curve of the hydrodynamic is steeper. To illustrate the limited capacity of mechanical friction brakes, as compared with these dynamic energy dissipators, Fig. 4 also shows the horsepower curves for two 54.7-in-diam × 10.5-in-face brakes, calculated on the basis of the thermal rating as given in this paper.

Fig. 5 more clearly illustrates the speed-horsepower characteristics of the two types of supplemental energy dissipators with respect to drawworks application, and also illustrates a basis for selecting a size to meet any general set of operating conditions. The example is for a drawworks having a 28-in-diam drum with the line unspooling from the second layer and for eight $1^1/_8$-in. lines strung through the blocks. Since the time scale is for total seconds required to run-in a stand, average values throughout the total elapsed time of running-in a stand must be used for the pipe speed scale and for the horsepower curves. The horsepower curves for the energy dissipators show their capacity for these average speeds. The horsepower curves for the suspended load are based upon calculated air weight of the loads as well as these average speeds.

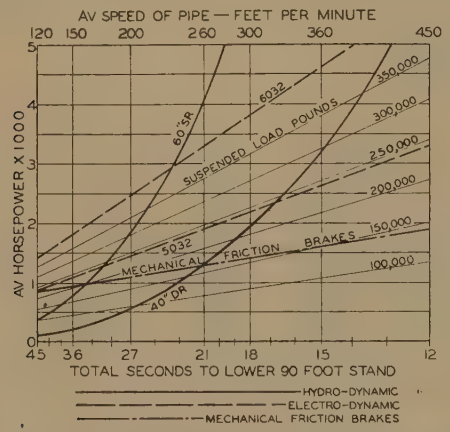


Fig. 5 Typical Performance Diagram of Supplemental Energy Dissipators and Mechanical Friction Brakes for Drawworks Application

[Drum diameter 28 in., eight $1^1/_8$-in. lines. Line unspooling from second layer ($31^3/_8$ in. diam).]

The actual rate at which horsepower must be absorbed by the energy dissipators varies widely throughout the running-in of a stand and is affected by highly variable factors of speed, friction, buoyancy of the fluid, etc. The averages of these actual horsepowers, if they could be determined accurately, would be less than the averages shown by the diagram. However, the actual horsepower capacity of the dissipators also varies with speed throughout the descent. These variable factors compensate for each other in this type of diagram, and field data and observations have shown that for any desired total elapsed time of running-in a stand, the intersection on the diagram of the time scale and the capacity curve for a given-size dissipator indicates satisfactorily the maximum load which should be handled with that dissipator. Also shown on this diagram is the thermal-horsepower capacity rating for two 54.7-in-diam × 10.5-in-face mechanical friction brakes, the average size furnished on drawworks which also would be equipped with one of this size dissipators.

Methods of Speed Control When Running-in Drill Pipe

The most practical method of speed control when running-in a stand is the one which, in consideration of the characteristics and limitations of the brake equipment, can be depended upon consistently to get the bit on bottom as speedily as is compatible with economy and safety.

Fig. 6 illustrates a preferred method of operation when the brake system permits its use. Although this figure is based on running-in a 90-ft stand in 15 sec, the curve is characteristic of runs made with other-length stands or in greater or shorter lengths of time. It will be noted that approximately 3 sec are required to accelerate the speed from rest at I to the maximum velocity at II. The rate of acceleration is slow at first, due to inertia and static friction, but after these are overcome, it becomes substantially uniform. The transition from the acceleration period to period of uniform velocity is rapid but smooth. By maintaining the maximum safe velocity for the conditions of operation during the 10 sec between II and III, the stand is lowered in the shortest possible time. In the final 2 sec, from III to IV, required to bring the load to rest, the deceleration starts rapidly but smoothly, and the final stop at IV is made without shock or jolt.

To consistently approximate this method of operation with mechanical friction brakes requires the close attention and alertness of the driller. It also requires that the brakes be maintained in first-class condition and not overloaded.

This method of operation may be closely approximated with the electrodynamic type of supplemental energy dissipator with a manual resistance control permitting the driller to vary the resistance rapidly as required during the acceleration and deceleration periods, the resistance being substantially constant during the period of uniform velocity from II to III. The mechanical brakes must be applied to bring the load to stop at IV.

This closely approximates the operation with a hydrodynamic type of supplemental energy dissipator which, without change of the resistance control during the entire time of running-in a stand, automatically permits the load to accelerate from I to II, and then automatically governs or holds the speed substantially constant from II to III. However, at III, the start of the deceleration period, the mechanical brake must be applied to finally bring the load to stop at IV.

Also shown in Fig. 6 is a method of operation more commonly followed when using mechanical friction brakes only, and by some operators when using electrodynamic supplemental energy dissipators.

When using this method, and to run-in a stand in the same 15 sec, requires that the speed must be accelerated to a value about 14 per cent higher than that for the first method described. The acceleration period from I to II is generally somewhat longer. The period of substantially uniform velocity, II to III, is generally shorter, and the deceleration period from III to IV is longer.

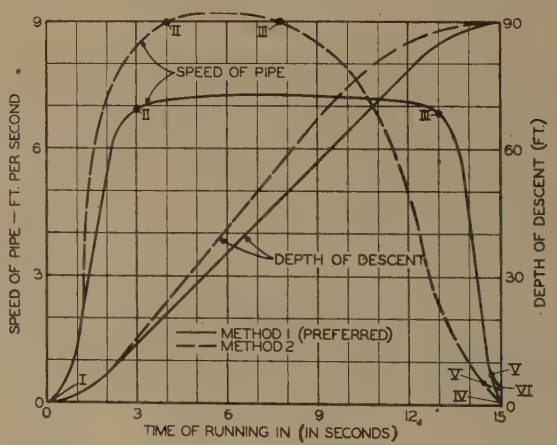


Fig. 6 Relationship Between Speed of Pipe and Time of Running-In a Stand for Typical Methods of Operation

When using mechanical friction brakes only, this method does not require as close attention on the part of the driller, and since the velocity and deceleration of the load are less just prior to the final stop at IV, it may be of advantage if the brakes are not in the best condition or tend to stick or grab.

It is a practical method of operation when using an electrodynamic type of supplemental energy dissipator with its quick-

acting resistance control and sufficient capacity at low speeds, so that the driller may quickly and positively vary the resistance as required throughout the descent.

When using this method with any type of supplemental dynamic energy dissipator, and unless some complicated form of automatic resistance control is also provided, any possibility of automatic speed governing is sacrificed for the reason that with any brake of any type and with a fixed setting of the resistance control, the energy-absorption capacity decreases as its speed decreases.

Another method of operation, which is a modification of the two just discussed, is shown by the deviation of the deceleration curves from V to VI. In this case the pipe is not brought to a complete stop by either the mechanical brakes or supplemental energy dissipator but is slowed down to a speed of approximately 0.3 to 0.5 fps at which the slips may be set.

This practice is followed by some drillers when using mechanical brakes only and running-in to shallow depths with light suspended loads. It is frequently followed when running-in to greater depths and using an electrodynamic type of energy dissipator which has relatively high resistance at very low speeds.

It is the author's opinion that the caution which must be taken in using this method does not justify the very slight saving in mechanical-friction-brake service otherwise required for actually stopping the load, and the possible saving of a small fraction of a second in time.

On the basis of brake service, even if the slips are set when the pipe speed is as great as 0.5 fps, the energy the brakes would otherwise have to absorb is not more than 0.1 per cent of the total dissipated by the brake system throughout the descent. With a 100,000-lb suspended weight, this is only 9000 ft-lb out of a total of 9,000,000 ft-lb developed during the descent.

On the other hand, this final amount of energy is not lost but must be absorbed by the rotary table and substructure. An instant of inattention or carelessness on the part of the driller in handling the controls, or failure of controls near the end of the deceleration period, may result in an accident, or the rig and pipe being subjected to severe and detrimental suddenly applied loads.

Relationship Between Rate at Which Energy Developed by Load Is Dissipated by Brake System

To evaluate completely the service to which drawworks brakes are subjected during the running-in of a stand of pipe, it is not only necessary to consider the method of operation but also, in so far as the variable factors permit, to analyze the relationship between the rate at which the energy developed by the load is dissipated by the brake system.

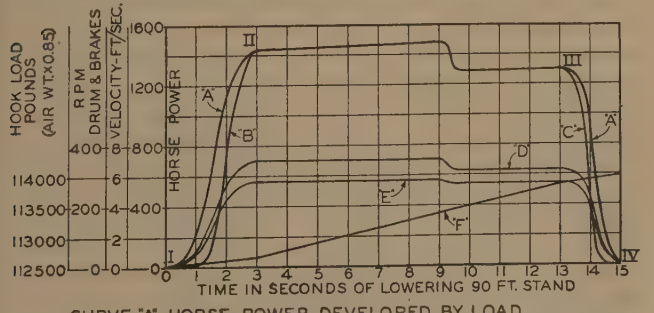


FIG. 7 TYPICAL RELATIONSHIP BETWEEN RATE AT WHICH HORSEPOWER DEVELOPED BY LOAD IS DISSIPATED BY DRAWWORKS BRAKE DURING RUNNING-IN OF A STAND OF PIPE

For the purpose of illustration, Fig. 7 shows the results of such an analysis based upon using a drawworks with 28-in-diam drum and 8 lines strung through the traveling block and lowering in 15 sec, by the preferred method of operation previously discussed, the 90-ft stand run-in when the bit is passing the 6000-ft depth. The calculated indicator weight of the suspended load at the start of the run has been taken as 112,500 lb. In addition to showing the velocity of the descending load, Fig. 7 also shows the rpm of the drum (and brake), the increase in suspended weight due to unspooling of the line, and the time rate at which the energy, expressed as horsepower, is developed by the load and dissipated by the brakes.

Although this example is for a specific set of conditions, the curves are typical and illustrate the general relationship which must exist between load, speed, energy developed, and energy dissipated at each instant during the periods of: (1) acceleration; (2) substantially constant velocity; and (3) deceleration, these three periods occurring in any method of operation.

Acceleration Period. For this example, 3 sec are required to accelerate the load from rest (at I) to a velocity of 7 fps (at II), and during this period the load is lowered 10 ft.

Five known and interrelating factors permit the values for the curves of the acceleration period shown in Fig. 7, and to an enlarged scale in Fig. 8, to be adjusted with relationship to each

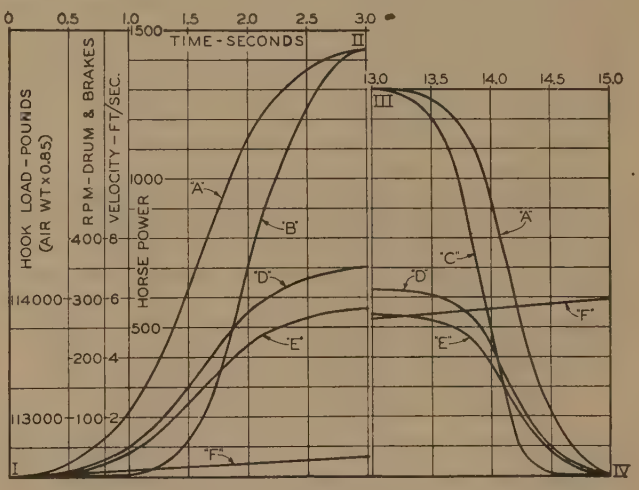


FIG. 8 (*left*) ACCELERATION PERIOD OF FIG. 7, TO ENLARGED SCALE

FIG. 9 (*right*) DECELERATION PERIOD OF FIG. 7, TO ENLARGED SCALE

other and plotted to a reasonably close degree of accuracy. These factors are as follows:

1 At any point in the descent of the load, the total energy, which up to then has been developed, is the product of the weight by the distance it has fallen. Thus at II, the total energy developed throughout the acceleration period is 112,500 × 10 = 1,125,000 ft-lb or approximately 10 per cent of the total for the entire descent.

2 The area under the "Hp Developed by the Load" curve during the acceleration period must represent the 1,125,000 ft-lb total energy developed throughout that period. For the curve as originally drawn with 1-in. vertical scale representing 200 hp (110,000 ft-lb per sec) and with 1-in. horizontal scale representing a velocity of 0.5 fps, the "scale conversion factor" for this relationship is 55,000 ft-lb per sq in., and the area under the curve to the original scales must be 20.4 sq in.

3 For any instant during the descent the horsepower being then developed by the hook load must equal (Hook load × ve-

locity in ft per sec)/550, and the values indicated by the curves for each instant must conform with this relationship.

4 Except at the very beginning and end of the acceleration period, the increase in acceleration, and corresponding increase in horsepower developed, should be substantially constant.

5 For any instant of descent, the value for the "Rpm of Drum and Brake" curve may be calculated from the "Velocity of Descent" curve.

It is obvious that since the acceleration period ends at II, at that point the 1,125,000 ft-lb total energy up to then developed by the load must become balanced by the sum of the energy stored in the system as a result of acceleration and that dissipated by the brakes and by friction of the rig. Also, for the velocity to become uniform at II, the rate at which horsepower is dissipated by the brakes must then equal the rate at which horsepower is developed by the load, less of course the friction. Consequently, the curves for horsepower developed by the load and horsepower dissipated by the brakes must meet and coincide at this point.

The "Hp Developed by the Brake" curve has been plotted for a hydrodynamic type of supplemental energy dissipator with resistance control set to develop the required horsepower at II. However, with some possible slight modification, it would apply for any other type of brake so operated and so controlled as to permit the load to accelerate to 7 fps in 3 sec. The area under this "Hp of the Brake" curve must represent the energy dissipated by the brake during the 3 sec of acceleration. By measuring this area and multiplying by the "scale conversion factor," it is determined this dissipated energy is 770,000 ft-lb or approximately 7.5 per cent of the total for the entire descent.

The difference in areas under the horsepower curves for the load and brake must represent the total of the potential energy stored in the system due to the acceleration of the moving weights and that lost through friction. The potential energy stored in the moving weights may be calculated accurately by the formula

$$(\text{Weight} \times \text{velocity}^2)/64.4 = (112{,}500 \times 7^2)/64.4 = 85{,}800 \text{ ft-lb}$$

Consequently, the energy dissipated by friction of the rig is

$$1{,}125{,}000 - 770{,}000 - 85{,}800 = 269{,}200 \text{ ft-lb}$$

The over-all efficiency of the rig during the acceleration period then becomes

$$100\,(770{,}000 + 85{,}800)/1{,}125{,}000 = 76.1 \text{ per cent}$$

which appears to be a reasonable value.

It is obvious from examination of these curves that for the first second or so of the acceleration period, practically all the energy developed by the suspended load is required to overcome inertia and friction, and the application of any substantial braking force during this time would extend materially the acceleration period. Any brake system having appreciable resistance at low speeds will require manual adjustment of control throughout the acceleration period.

Period of Substantially Constant Velocity. It is not necessary to make an analysis of the energy balance between load and brakes during this period (from II to III) because, since the speed is held substantially constant, the energy developed by the load, less the friction of the system, must be dissipated by the brake system at substantially the same rate. Any slight acceleration or deceleration during the period from II to III would only cause a slight increase or decrease in the potential energy stored in the moving weights. The dip in the horsepower and velocity curves of Fig. 7 at 9 sec is the result of the line starting to spool off a lower layer.

However, it is of interest to note that during this period the energy is dissipated at a maximum sustained rate, and, if the brake resistance balances this rate, no regulation should be required of its resistance control. Also, that during this period approximately 79 per cent of the total energy developed during the entire descent is dissipated.

Deceleration Period. During the deceleration period the same five known and interrelating factors apply as for the acceleration period, and, in this example, for the 2 sec taken as the time to decelerate the load (from III) to rest (at IV), the load is lowered the last remaining 10 ft.

The total energy remaining in the system at III, and which must be completely dissipated by the time the load is brought to rest, can be calculated definitely by the formula

$$(\text{Suspended weight} \times \text{velocity}^2)/64.4 + (\text{Suspended weight} \times \text{distance to be lowered})$$

$$(113{,}820 \times 6.28^2)/64.4 + (113{,}820 \times 10) = 1{,}207{,}900 \text{ ft-lb}$$

Assuming the same over-all efficiency of 76 per cent applies to the deceleration period as for the acceleration period, the remaining energy that must be dissipated by the brake system becomes $1{,}207{,}900 \times 0.76 = 918{,}000$ ft-lb, or about 9 per cent of the total developed during the entire descent. The area under the Hp Developed by the Load curve also shown to enlarged scale in Fig. 9, and which represents this energy, may be determined by using the 55,000-ft-lb per sq in. scale conversion factor. The curves for "Hp of the Load," "Velocity of Descent of Pipe," and "Rpm of Drum and Brake" may now be adjusted and plotted in the same manner as for the acceleration period.

This Hp of the Load curve also represents the time rate of energy dissipation by whatever brake system is used during the deceleration period, and averages 834 hp over the short period of but 2 sec, as compared with an average rate of 1400 hp over the sustained 10 sec of the constant-velocity period.

If mechanical brakes only are used during the deceleration period the 918,000 ft-lb total energy (834 hp average) to be dissipated in 2 sec must all be absorbed by them.

If an electrodynamic energy dissipator is used and the resistance is so controlled that the mechanical brakes are required to do no work up to the instant of stop, then these values represent the total energy of which the mechanical brakes are completely relieved during the deceleration period.

To illustrate the result when using a hydrodynamic energy dissipator, its horsepower capacity, during the deceleration period, plotted on the basis of its speed, and with the same setting of resistance control as used to govern the speed up to point III, is shown by C, the "Hp Developed by Hydrodynamic Supplemental Dissipator" curve. The energy dissipated by it during the deceleration period, represented by the area under this curve, is determined to be 646,250 ft-lb. The remaining energy then required to be dissipated by the mechanical friction brakes is represented by the area between the two horsepower curves and becomes 918,000 — 646,250 or 271,750 ft-lb. This is only 28.5 per cent of the total developed during the deceleration period, and only 2.67 per cent of the total developed during the entire descent. Under these conditions, and if only 8 gal of 80-deg cooling water were circulated per minute, the operating temperature of the rims should remain below 100 F. Even at 12,000 ft the operating temperature of the rims should remain below 120 F with only 8 gal of 80-deg cooling water circulated.

Methods of Driving Low-Speed Supplemental Energy Dissipators

The most common method of driving supplemental energy dissipators from the drawworks shaft is through a flexible cou-

pling capable of compensating for any slight misalignment of assembly.

With the electrodynamic type of energy dissipator, the manufacturer recommends, for reasons previously discussed, that a plain flexible coupling be used and that the dissipator be allowed to rotate at all times.

With the hydrodynamic type of energy dissipator, a combination flexible-cutoff coupling may be used to advantage, particularly with the larger sizes. By declutching when the braking action is not required, any inertia effects of the rotor are eliminated from other operations of the drawworks, and shaft seals and bearings are relieved of any unnecessary wear. The shaft-drive assembly is also relieved of any reversal of stresses caused by inertia, and the frequent changes in direction of rotation when running-in or coming out of the hole.

An overrunning or free-wheeling type of clutch is of particular and further advantage for use with the hydrodynamic type of energy dissipator. This clutch automatically and positively engages the energy dissipator at the instant the drawworks drum starts rotating in the direction to run-in a stand, and the energy dissipator is not revolved when raising the blocks. It eliminates any possibility of the drive shaft and keys being subjected to shock-loading caused by delayed clutch engagement at the start of running-in a stand or, more seriously, by the clutch engagement after the drum has attained a high speed.

High-Speed Supplemental Energy Dissipators and Their Drives

In recent years the use of small high-speed supplemental energy dissipators, driven from the drum shaft through a chain speed increaser, has become increasingly popular, the only practical limitations to their use being the safe maximum operating speed of the dissipator, and the safe maximum operating speed and load capacity of the chain drive.

To date the most common installations of high-speed dissipators have been on drawworks designed for operation to 6000 ft or less. They are also used on some drawworks recommended for 7500 to 8000 ft and have been adopted for one drawworks recently introduced and recommended for 15,000-ft drilling.

Advantages of high-speed energy-dissipator installations include compactness of assembly and greater flexibility as to location on the drawworks, as well as the possibility of faster operations owing to their having a considerably lower inertia factor and less drag for reverse direction of rotation. In the case of hydrodynamic energy dissipators, the small high-speed types have the further advantages that they lend themselves to simpler circulating-water arrangements, more flexible resistance regulation, and quicker response. This is due to the fact that both high and low-speed dissipators require the same total water circulation to dissipate the same amount of energy but, with the high-speed dissipator, the circulation is faster and the volume of water in the case at any time is less.

An obvious disadvantage is that the chain drive represents an additional element requiring some attention and eventual replacement. It is also true some types and designs of drawworks lend themselves more readily and practically to the slower-speed direct-driven dissipators.

In selecting the size dissipator and chain-speed ratio for any particular application, the maximum horsepower capacity required at minimum operating speed should first be determined. This will establish the required speed ratio and assure ample capacity for all conditions of operation. The maximum speed of the dissipator and chain for the top speed of drawworks operation should then be determined and checked against that for which such equipment is recommended.

In most drawworks installations the chain speed-increaser drive has a ratio of from 4 to 1 to 5 to 1. With a maximum drawworks drum speed of 400 to 500 rpm, the maximum speed of the dynamic energy dissipator may be about 2000 rpm, and the maximum chain speeds may be in the order of 4000 fpm.

Practical considerations and space limitations dictate that the chain drive be selected on the basis of a higher permissible load rating than that normally used for other chain drives of the drawworks. Such a higher rating, however, and as has been proved by actual field experience with drawworks, sand-reel, bull-wheel, mine-hoist, and other installations, is justified by the speed-resistance characteristics of the dissipators and the usual conditions of operation of the application.

In determining the load basis on which the chain drive should be selected, and these comments also apply to the selection of a coupling used with any type of drawworks installation, consideration must be given to the important fact that the load is increased as the depth increases and as each individual stand is run-in. Usually the maximum load is imposed only when lowering the last stand of pipe to the maximum depth to which the well is drilled.

Consideration must also be given to the fact that while the total of all the time expended in running-in pipe during the drilling of a well varies under different drilling conditions, it is a small percentage of the actual total drilling time and that the brakes are actually used for only about 25 per cent of this total running-in time. Another favorable factor for the drive is that the energy dissipator with its control must be such as to definitely assure a smooth even application and increase of braking force, and the impossibility of a suddenly applied load on the chain at any time during the braking operation.

For the more general classes of service for which chain drives are used, the load-carrying capacity of the chain is usually determined by applying an application factor of 15 to 25 to its ultimate strength. The selection of the chain for each dynamic-energy-dissipator application to drawworks should be made on the basis of its own design requirements and the recommendations of the chain manufacturers. However, in view of the favorable conditions of the application as described, application factors as low as 10 or less, based upon the maximum horsepower anticipated for the running-in period, in some instances have been recommended by chain manufacturers and used with satisfactory field results.

Automatic Drilling Feed Controls for Rotary Drilling Rigs

By W. S. CRAKE,[1] HOUSTON, TEXAS

A considerable amount of detailed study and thought, design, and field testing of various means of achieving the desired results has gone to bring work on drilling controls to their present status. Up to the present, however, the results of this study and work have not been recorded except in improvement in machine performance. This paper is not primarily intended to be an evaluation of the results of automatic-controlled versus hand-controlled drilling, but it is felt by the author that the subject is of major importance in reducing drilling costs, and that all those factors which have been considered, and the analyses of the problems, should be placed on record. Others may wish to extend or revise the analyses and build or purchase similar devices, and this report will, it is hoped, save them from having to duplicate the efforts of the past few years and to determine if the machines designed or offered have the required characteristics.

IT has long been recognized that the maintaining of constant loading on drilling bits will result in more footage per bit and more footage per unit of time, all other factors being equal. The result is fewer round trips with their horsepower-hours, wear and tear on rigs, and so forth.

The best hand-feed brake lever is a poor substitute for full automatic control, since worn brakes creep off and either "drill off" or "pile on" weight. Most are worse than this in that frequently they drop off from 1/4 to 1/2 a drum revolution suddenly when worn. It is physically impossible for a man to feed constantly for long periods, just as it would be on a lathe or other machine tool.

The author intends to avoid the term "automatic driller" in case any reader should misconstrue the purpose of automatic drilling controls. They make the driller more indispensable than ever, but require him to exercise his judgment more than his muscle. There is no insinuation that a constant single weight is desirable for drilling any more than a constant machine-tool spindle speed or feed rate, or a single-type bit, would drill a composite block of plates of lead, steel, copper, concrete, and clay satisfactorily.

With automatic control the driller can watch his drilling-rate chart and his other drilling symptoms and adjust his feed weight and rotating speeds to obtain the best possible results. By making constant one of his drilling variable, i.e., the weight on the bit, the rotary speed automatically steadies down in any given formation and the table behavior becomes a very sensitive indicator of formation changes and types. The penetration-rate indicator, or recorder, in most cases gives a clear graphic story of a bit becoming dull and can save many additional hours of "hoping that the bit will make a few feet more."

[1] Chief Mechanical Engineer, Houston Area, Shell Oil Company, Inc.

Contributed by the Petroleum Division and presented at the Petroleum Mechanical Engineering Conference, Amarillo, Texas, October 3–6, 1948, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—PET-4.

ANALYSIS OF THE PROBLEM

Relationship of Bit Load to Hook Load. The first question is "what is the relationship of the hook load to the bit load?" If there is not reasonable similarity, the problem becomes complicated, or the solution impractical. The measured load on the hook is a relatively simple starting point, but measuring the load on the bit at the well bottom requires a vast amount of research and equipment yet undeveloped. The cost and complexity of this work could be out of proportion to the gains to be made.

Synthesis of the System. In order to arrive at an answer to the foregoing question, let us synthesize the suspended system into analogous parts. Referring to Fig. 1, the drill pipe string can be assumed to be a simple spring of length L, and the drill collar assembly becomes a short, stiff, heavy spring which, in effect, is a weight on the end of the drill pipe.

In Fig. 1 it is assumed that, with customary good drilling practice, enough drill collars are used to maintain the "neutral point" of load change from tension to compression within the drill-collar assembly. This assumption is not necessary to the analysis or correct usage, but is used in the example.

Referring to Fig. 1

W = weight on hook
W_1 = weight of drill pipe
W_2 = weight of drill-collar assembly
W_3 = weight on bit
W_4 = remaining weight of drill collar suspended on drill pipe

Since all wells are to some extent crooked, all the wall contact points, either steady, or periodical caused by vibration, cause friction forces between drill pipe and wall. Then, in order to start the whole string in motion, without rotation, off bottom

$$W = W_1 + W_2 - \Sigma f \quad [1]$$

and, on bottom

$$W = W_1 + W_4 - (W_3 + \Sigma f) \quad [2]$$

If the pipe is lowered slowly off bottom, as while drilling, the weight indicator will read the result of Equation [1] and, if on bottom, the result of Equation [2].

If, however, the pump is started, thus vibrating the pipe, or if the pipe is rotated, the pipe weight plus W_2 or W_4 wipes out the friction effect for all practical purposes, except as a damping action against longitudinal vibrations. The weight will then read

$$W = W_1 + W_2 \text{ (off bottom)}$$

and

$$W = W_1 + W_4 \text{ (on bottom)}$$

These conditions are analogous to the use of a vibrator to destroy friction in the movement on pressure-gage testing, or to the practice of rotating the piston of a deadweight tester for the same purpose (see Appendix).

By studying Fig. 2 the value of f_n, one of the numerous friction forces caused by, say, a tool joint touching on the wall, becomes $\mu a W_x \sin \theta$, and the sum of all these friction forces equals $\Sigma \mu a W_x \sin \theta$, or Σf_n, where μ = coefficient of friction.

As long as W_x is greater than f_n, the foregoing reasoning holds. If an extremely crooked hole or bent pipe causes W_x to be less than f_n, then this latter amount becomes a reduction in hook weight.

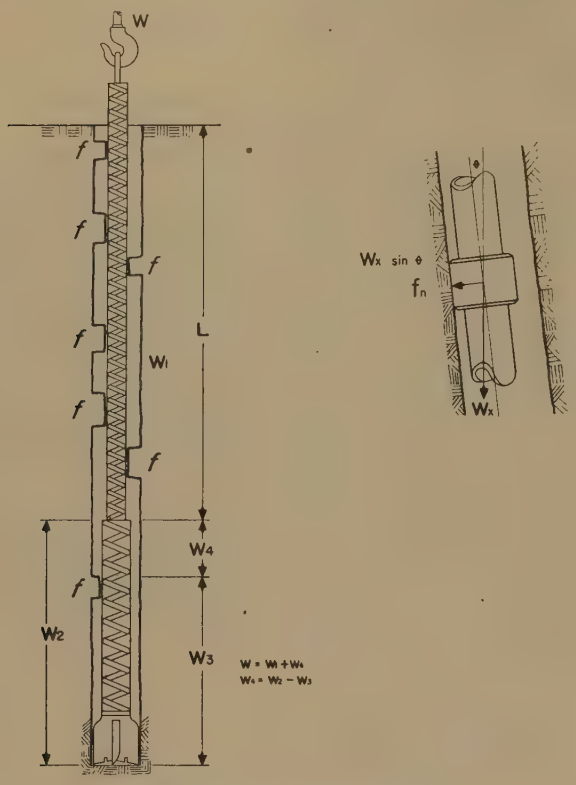


FIG. 1 (*left*) DRILL-PIPE STRING

FIG. 2 (*right above*) FRICTION FORCE AT ANY CONTACT POINT (k = friction coefficient with mud lubrication; A = area of contact; W_x = suspended load; θ = angle of deflection from vertical at this point.)

While rotating, Σf_n becomes torsional friction, which remains as a torsional or horsepower, load on the rotary table.

The effect of torsion on the length of the drill-pipe string is shown in Appendix I.

Strain in System. Since the whole system is elastic, the strain (or sketch) of the pipe is obtained from the formula

$$e = \frac{72\,L^2}{E}\,[\delta - 2w(1 - m)] \quad \ldots\ldots\ldots\ldots [3]^2$$

where e = strain or stretch, in.
L = length of string, ft
E = Young's modulus, 30,000,000 psi
δ = density of steel, lb per cu in.
w = density of flotant, lb per cu in.
m = Poisson's ratio, or 0.28

The effect of adding drill-collar weight W_2, is to add an equivalent load, less the buoyancy factor, in feet of pipe of a given weight. The equivalent length of pipe, with drill collars suspended, equals

$$\frac{W_2\left(1 - \frac{w}{\delta}\right)}{U} = L_2$$

where U is the weight per foot of material in string L.

[2] Spang-Chalfant Engineering Data Sheet 10-050.

On bottom, the equivalent length of pipe with W_4 still suspended, equals

$$\frac{W_4\left(1 - \frac{w}{\delta}\right)}{U} = L_3$$

By recalculating the strain in Equation [3], using $L + L_2$ or $L + L_3$ in place of L, new strains become easily determined. The difference of strain between the two readings just given becomes the distance the kelly must lower after first contacting bottom, in order to place load W_3 on the bit. These may be obtained more easily by reading the strains from the handbook.[2]

The existence of these strains actually provides additional equivalent time lag to help the controller function smoothly.

Forces Acting on String, and Their Effect on Hook Load. Forces acting on the string when "feeding off" are listed in the following (all forces acting downward are listed as positive, and upward-acting forces as negative):

1 Weight of equivalent pipe string (as mentioned) ..... +
2 Wall friction, depending on direction of movement + or —
3 Fluid friction of circulating fluid in drill pipe ..... +
4 Friction drag of return fluid on OD of drill pipe ..... —
5 Force of closed-in fluid at upper end of hose ..... —
6 Force of semiclosed-in fluid pressure at bit (pressure drop across bit times inside-diameter area of pipe) ..... +
7 Reaction effects of jets in bit, including bit fouling, etc. . —
8 Torsional stress on drill pipe ..... none
9 Vibratory effects due to critical speeds, kinked pipe, "rough bottom" (usually of some harmonic order + and — damped by other forces) ..... depends on circumstances
10 Friction through kelly bushing ..... —

When off bottom, rotating as slowly as possible without circulation, the author feels that (1) is positive and is close to weight of the portion of the load not immersed plus (weight of immersed portion times mean buoyancy factor); (2) is zero, having been converted to horsepower loss; (3), (4), (5), (6), and (7) are zero; (8) is a low figure of horsepower into rotary table and has no effect on hook load; (9) is zero; (10) is zero, no vertical motion is involved.

When off bottom and rotating at drilling speed with the pump at drilling speed (1) is as before; (2) is negligible (see Appendix); (3), (4), (5), (6), and (7) are of the sign shown and indicate their algebraic sum on the weight indicator; (8) has no effect on hook load; (9) these may show up through the damping effect of wall friction and fluid friction: If so, they are self-measuring and, if serious, are visible at the kelly or audible on the rotary table; (10) cancels out unless blocks and hook are swaying, or are of low order if the kelly is running true in the hole and bushings are of antifriction type. In any case, this force becomes one of the f_n forces under Σf as previously mentioned.

On bottom, just as the bit contacts, all the factors given are measured automatically and give a net compensated reading on the weight-indicator instrument. As long as the rotary table turns, items (2) and (10) remain practically constant and, as long as the pump runs at reasonably the same speed, items (3) and (7) remain practically constant.

Vibrations caused on the derrick, blocks, and pipe by pump pulsation are external vibrations which can be removed by damping the pump or the instruments, to eliminate them. They are not, it is believed, transmitted to the bit unless they are so large as to be almost dangerous to the surface operations. The "elastic-bag" nature of the drill pipe, friction on walls, and immersion in fluid should give almost constant fluid flow, at the lower

end of a long drill-pipe string. In addition, the restriction at the bit jets gives further stabilization to the flow rate.

Vibration of cutting action, such as the "washboard effect" of raising and lowering the drill collars each time a rock-bit tooth "goes over center," is absorbed in the drilling string, unless the bottom of the hole becomes badly out of level.

In view of these factors, it is seen that a reading of the hook load becomes one in which all the major forces on the string become self-balancing under reasonably steady-state conditions, and, it appears, over a relatively wide range of variation of forces. The major forces to consider, therefore, become the off-bottom summation, and the on-bottom summation of those factors which cause variations, such as wall friction, mud-flow rates, rotary speeds, and vibration.

The difference shown on an accurate indicator is, subject to damping out unusual and abnormal forces, very close to the true mean weight of the bit.

Instantaneous bit loads, undoubtedly, are quite wide in their variations and, if we could obtain a reading of these forces at the surface, it is believed they would have to be damped to about the same order as the surface readings are naturally read, before being usable as a controller signal.

Furthermore, to those who consider bottom-hole measurement of bit loads more desirable, the author wishes to point out that, after damping these signals to usable form, the interpreter, or instrument, has to apply the result to the machinery at the surface in any case. Therefore if a bottom-hole controlled variable was used, the entire elastic string would have to be moved to correct an inelastically obtained reading.

In other words, by using the hook signal and controlling the hook load, the effect of position changes on the string elasticity comes in before pickup and after controller action and not between them. This reduces time lag between signal, control action, and resignal. The effect of string elasticity follows, as a natural damping action, and tends to reduce hunting of controllers to a minimum.

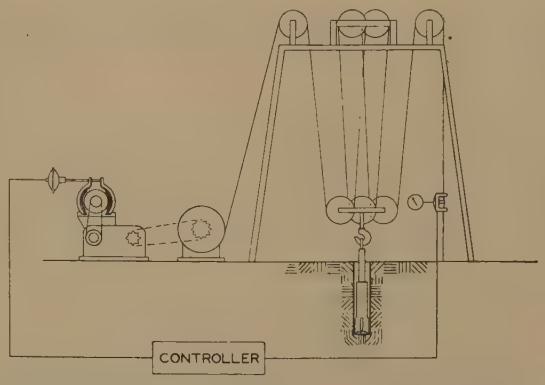


FIG. 3 ELEMENTARY DRILLING RIG, WEIGHT CONTROL

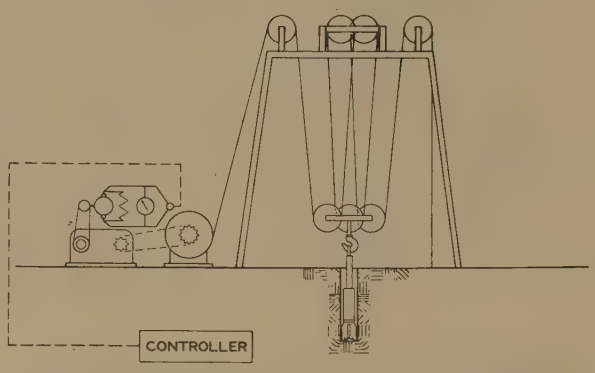


FIG. 4 ELEMENTARY DRILLING RIG, SPEED CONTROL

ELEMENTS OF CONTROL SYSTEMS

Pickup or Measuring Devices (Primary Elements). In a drilling controller, an elementary form of which is shown in Fig. 3, a signal, or controlled variable, proportional to the weight carried on the hook is necessary as the input to any weight controller. Such devices may be installed on the dead line and measure hydraulic pressure, as with a Martin Decker weight indicator, by air pressure, extension of a bar, or strain, using magnetic or resistance methods. Summations of strains on derrick legs and cables or any other means of obtaining a pressure or electrically controlled variable proportional to cable load are acceptable. The devices may be used on the live-line or drum end, including torque-sensitive devices, hydraulically or pneumatically supported torque arms, or hydraulic pumps, the discharge pressure of which is maintained at a set figure to represent torque load. All these devices should be compensated fully for temperature variations caused by operating or atmospheric conditions.

If a speed-limiting device is required or incorporated in the control, a second controlled variable, proportional to hook-lowering speed, is required. This may be a flyball governor, a tachometer generator, a hydraulic pump or air compressor, from which the output volume is controlled, and the output pressure then becomes proportional to the speed of operation or drum rpm.

Time Delay or Transmission. In order to enable controller mechanisms to perform their functions properly, especially at the higher drilling rates, sufficient time must be given for the load signals to reach the controller, actuate it, pass the result to the braking device, and make the motion adjustments necessary to correct the load back to its set point. This delay must also give time for overcoming inertia of the system and bringing the heavy

TABLE 1 PICKUP OR PRIMARY SIGNAL DEVICES[a] FOR AUTOMATIC DRILLING CONTROLS

Type of signal	Electrical	Pneumatic	Hydraulic
Weight, or load	Resistance strain gage Magnetic strain gage	Dead-line angular displacement—Thrustorq type	Dead-line angular displacement (Martin-Decker type)
	Dead-line angular displacement against hydraulic or pneumatic device	Torque arms with null balance against load	Dead-line full load on hydraulic cylinder
	Torque-measuring devices		Torque arms and null balance against load
	Null-balancing against main-drum torque		
Speed, or feed rate	Tachometer generator	Small compressor discharge against variable orifice. Pressure across orifice is proportional to speed	Pump discharge against variable orifice. Pressure across orifice is proportional to speed
	Requires a speed indicator such as an electric tachometer to give true speed	Requires a speed indicator such as an electric tachometer to give true speed	

[a] All devices must be fully temperature-compensated, or unaffected by temperature, over operating range.

inertia of drilling-rig moving parts back through the transmission, blocks, and hook to a new speed or weight setting. The final result eventually gets back through this system to the dead line and starts all over again.

If the time constant of the moving-load parts, or that time taken by those parts to accelerate due to a change in load, is less than that of the total cycle time of instrument action, hunting becomes uncontrollable and stability is impossible.

On a drilling rig, it takes a definite time for surface positional changes to travel the length of the drilling string and for the string to adjust itself to its new position. This time depends on the rate of drilling, rotational rate, friction forces, and depth of well.

A control system with too fast action would probably wear itself out rapidly by attempts of the controller and brake to adjust to its signal variations. For this reason, a time delay, in the form of a speed increaser from the drum shaft, is highly desirable, if not essential to a good control system. It gives elasticity and flexibility which, under the widely varying operating conditions on an oil well, are essential.

Brake. Any form of brake used, whether electrical, mechanical, pneumatic, or hydraulic, must be of a type sensitive to the slightest motion of the controlling mechanism. For this reason, variations in friction characteristics over small ranges of adjustment are highly undesirable. Since the brakes are not of the positive-positioning type but should "float on the load," variations of characteristics due to temperature or of friction coefficients, except as just noted, are not important. The exception to this is that the torque required to overcome static friction of the braking mechanism should be kept to an absolute minimum. This is because the brake often shuts the load down to a static position and the sudden overcoming of static friction always causes an overrun or hunt. Any eccentricity or lost motion in the brake mechanism is extremely bad in such a case. If mechanical friction brakes are used, it is recommended that they operate oil-immersed, to minimize the problem described.

Control Instruments. Figs. 5 and 6 show simple diagrammatic forms of drilling control systems. The system may be described as follows:

The signal or controlled variable is connected to the controller. As a rule, the transmission lag is of very low order. The controller must have the control point adjustable over a wide range.

Since the brake or retarding device is floating in a delicately balanced condition against the load, a sensitive proportional-band adjustment is required in the instrument. This proportional-band adjustment is sometimes known as a sensitivity adjustment, and determines the rate at which brake force is applied per unit of controlled variable change. Since drilling-rig con-

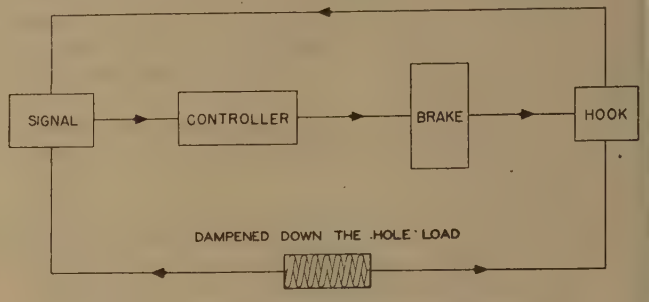


FIG. 5 DIAGRAMMATIC WEIGHT-CONTROL SYSTEM

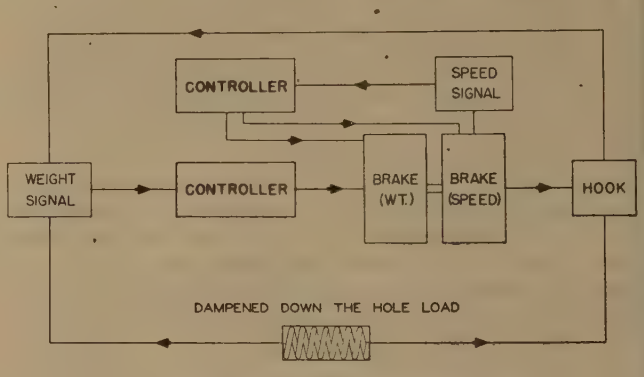


FIG. 6 DIAGRAMMATIC SPEED AND WEIGHT-CONTROL SYSTEM

TABLE 2 TYPES OF BRAKES FOR AUTOMATIC DRILLING CONTROLS

Electrical	Pneumatic	Hydraulic	Mechanical
Motor balanced against torque load. Can be over-excited to raise load	Air compressor operated against variable orifice. Air pressure on pistons holds load. Controller operates motor valve on discharge of compressor	Oil pump operated against variable orifice to control pump pressure. Fluid pressure on pistons holds load. Controller operates valve on discharge of pump	Precision friction brakes, preferably oil-immersed. Brake applied by motor valve, hydraulically or pneumatically operated
Eddy-current brake			
Balanced torque-arm type with motor to drive feed-gear train	Balanced torque arm, with air motors to drive feed-gear train	Balanced torque arm, with fluid bleed to lower load—dual brakes used (Kinzbach)	
		Hydraulic variable-speed transmission using constant-torque, variable-stroke pump. Speed governs stroke. Weight governs pressure	

TABLE 3 CONTROLLER TYPES FOR DRILLING CONTROLS

	Electrical	Pneumatic	Hydraulic
Weight	For use with brakes. Null-balance control of armature current of brake motor. Null balance-control of field current of eddy-current brake. Electronic controls give highest speed of operation, great flexibility	Pressure controller, using motor valve to operate brake. Signal from weight devices, Table 1. With floating proportional-band adjustment. Fast and flexible operation	Hydraulic pressure controller with brake-operating motor. May employ signal from any source, Table 1. Faster but less flexible than pneumatic. Requires special damping
Speed	With electric brakes. Control of field voltage of brake motor from tachometer-generator signal	Pressure controller as above, with motor valve to operate brake. Signal from hydraulic or pneumatic devices, Table 1	Pressure controller similar to above, with motor valve to operate brake. Signal from any source, Table 1
	Control of eddy-current brake field from tachometer generator signal. Superimposed on, or in series with, weight control		

trollers are close-coupled to the braking means, the controller lag, or time taken for a brake response to the controller, is small.

The action of the control is as follows: As the bit drills off, the controlled variable changes and operates the controller which, in turn, releases the brake. The drill-pipe position moves downward until the bit weight goes back to its required value and thus returns the controlled variable to its original value. At this time it again operates the controller to tighten the brake. This cycle, continually repeated, accomplishes a balance on the system. Actually, the changes take place so rapidly and proportionately that smooth pay-off is obtained when the system is designed properly.

Due to the wide range of weights and speeds to be controlled, a proportional-band adjustment is necessary in the control instrument in order to obtain stability of control at various ranges of set point.

Simple standard forms of instruments, such as pressure controllers, are readily adaptable to drilling controls. In view of the small percentage of instrument reading which is, in reality, being controlled, devices which amplify the magnitude of variations of controlled variables are desirable. Ratio-type relays or pilots can be used to multiply the output of the controller per unit of change of controlled variable. These devices, however, multiply the controller errors and lost motion, and it is more desirable to

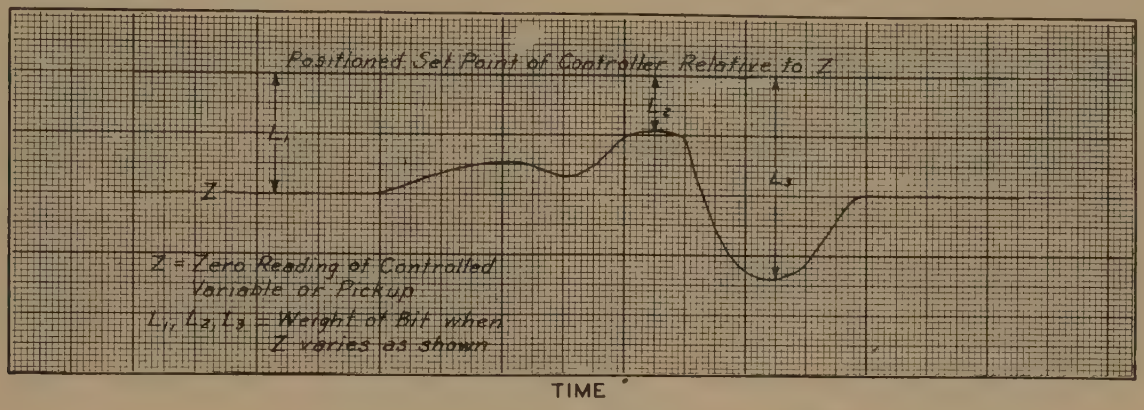


FIG. 7 EFFECT OF PICKUP DEVICE ZERO ERROR ON BIT WEIGHT

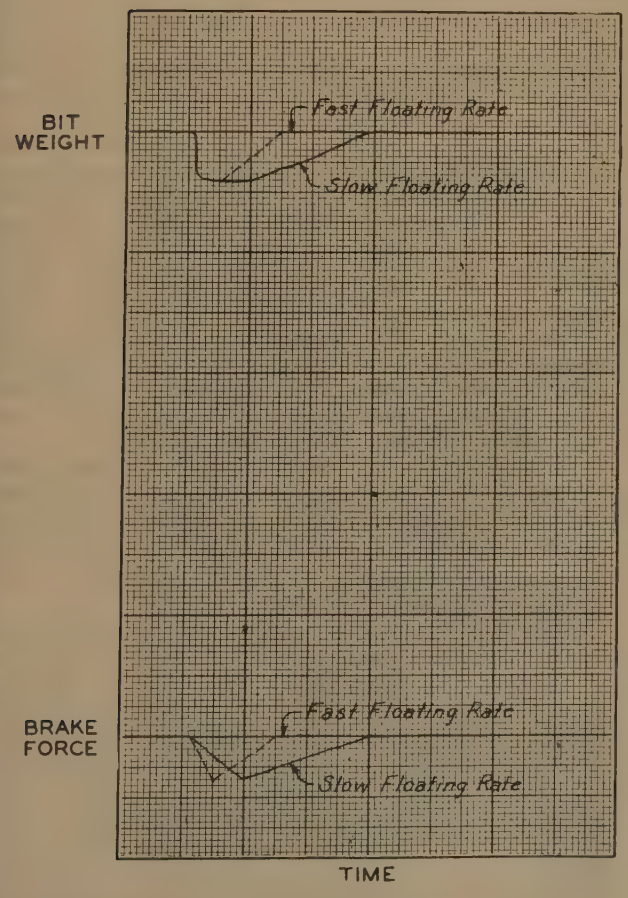


FIG. 8(*a*) RESPONSE OF PROPORTIONAL SPEED FLOATING CONTROLLER TO SUDDEN CHANGES

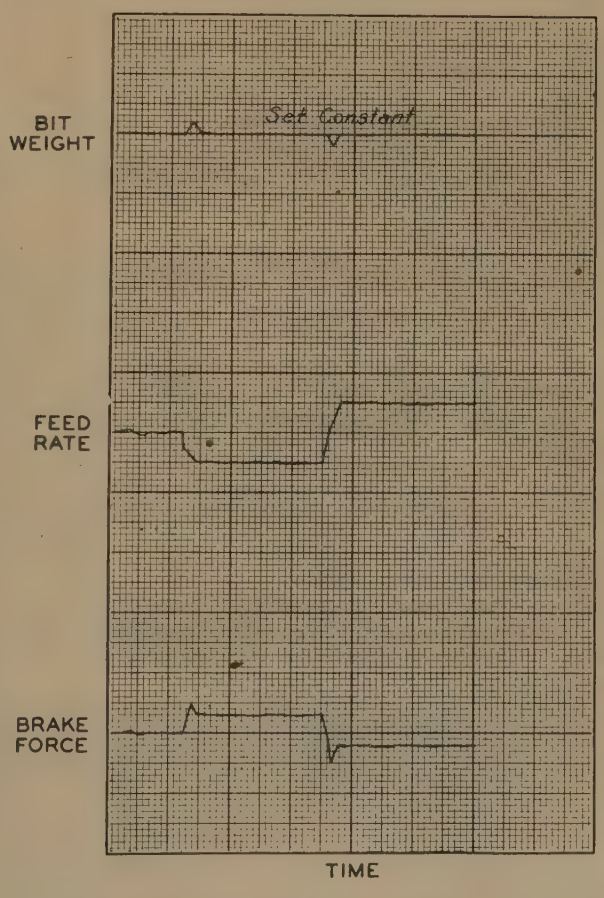


FIG. 8(*b*) RESPONSE OF PROPORTIONAL SPEED FLOATING CONTROLLER TO SUDDENLY APPLIED SUSTAINED CHANGES; FAST FLOATING RATE

multiply the controlled variable changes ahead of the controller if these changes are not already of the required order.

As with any instruments, properly engineered installation and reliable power supply must be furnished. This means, for example, that clean dry air must be furnished. Instruments may be complex, but the operating adjustments must be simple so that the drilling crews can use them properly.

Horsepower, Range, Sensitivity, and Accuracy

"Pickup" or Primary Element. This device must be self-operating from the source of power available, i.e., the load. On heavy drilling rigs it should have sufficient capacity to transmit the maximum load expected, without rupture; usually involving impact loads on cables much heavier than the normal loads. For instance, a cable-operated device should be capable of having the cable broken without causing mechanical failure or, if possible, without permanently affecting calibration of the device.

On rigs using eight lines and for hook loads up to 360,000 lb, the operating-load range may be assumed to be about 30,000 to 360,000 lb, which is, neglecting friction, from 3750 to 45,000 lb on the dead-line cable. Pulls up to cable braking strength need not register.

The device should have a minimum accuracy over full operating scale of ±1 per cent and, when the load is full floating, or subjected to rig operating vibrations, should have a sensitivity, or be able to transmit changes in hook load, of 150 lb or less. This is about one part in 2500 or 0.04 per cent.

Controller. The controller should be of a precision type and, for general oil-field use, should not incorporate copper or brass moving parts which are affected by sulphur gases. Stainless parts are most desirable, but such parts as are necessarily of corrodible nature should be protected by coatings capable of outdoor exposure in oil-field vapor.

Temperatures of from 150 F in the sun to —50 F in the winter should not affect instrument operation at any given set point, providing proper "winter" or "summer" treatment of the instrument is specified and used. Air supply, particularly, should be dried properly so that freezing of pneumatically operated parts will not take place in winter.

The control system, as a unit, must be able to recognize the change of 1:2500 from setting point which the primary element shows and cause the corresponding correction force to be applied to the brake. Since, as described before, the whole system is balanced delicately, this presents no serious problem with standard precision equipment, because, in the delicate state of balance, the slightest change in opposing force causes a change in braking force. In case of a pneumatically operated system, the valve which operates the brake has to move only a few thousandths of an inch to change the force materially. Brake-force change therefore involves practically no increased air-volume space to be filled. Compression volume is practically the whole required instrument output.

On hydraulically operated mechanical brakes, the foregoing is a detriment and can cause violent hunting unless a cushion chamber or other damping device is incorporated to smooth out change in the rigid hydraulic operating medium.

The range of instruments will depend on the value and type of signal from the primary element and, also, the medium and value thereof, used for transmitting the controller output to the brake. A wide control-point setting range, of about 12:1, is desirable.

Accurate knowledge and specification of the problem are always essential and will enable proper design and selection of instruments to be made.

Brake. The brake may be of various types i.e., electric, hydraulic, mechanical, or pneumatic.

In regard to capacity, it may be assumed that any feed-off device should be capable of feeding at the maximum drilling rate expected in the area to be drilled. In hard-rock territory, a speed range of from 0–60 fph would cover all except very soft formations. In soft-drilling territory, a rate of at least 0–120 fph should be covered by the controller. A recent paper by Nolley, Cannon, and Ragland[3] indicates that a maximum speed of 300 fph may be necessary to obtain maximum drilling rates with drag bits. This would raise the speed or control range to 1200:1 instead of 480:1. In order to attain this range, it is probable that brakes should have two ranges of speed, say, 0–60 fph for "hard-rock drilling," and 0–300 fph for drag bits. This, in the same unit, would probably mean a mechanical speed change, and the brake would absorb about three times the horsepower at the high operating speed. The load to be handled, and consequently, the horsepower to be absorbed, depend on the hook load. On ultra-heavy rigs, the maximum drill-pipe hook load may reach 350,000 lb, and the power, at 100 per cent efficiency and 120 fph, would be 21.22 hp. A light rig for 6000 feet drilling with 4½-in. drill pipe should be able to handle 125,000 lb on the hook, or 7.58 hp at 120 fph.

The brake and cooling system must be capable of radiating or absorbing the heat equivalent of the horsepower given.

In order to obtain sensitive and proportional effect to the action of the controller, the brake must be of an accurate precision type. All eccentricity or inaccuracy will tend to cause poor brake operation.

On mechanical brakes, up to the present time, speed increasers of about 800:1 ratio have been used between the drum and brake shaft but with development of proper controls it is believed this can be reduced to one half this ratio. Until it can be reduced to about 300:1, however, at least a quadruple reduction transmission from the drum shaft to the brake will be required. These transmissions provide the time lag between the drum and the brake to give instruments time to operate and control.

Causes of Errors. In the mechanical transmission, friction should be kept to a minimum. This is largely because, once the brake stops, quite a difference in hook loading, perhaps 1000 lb on a poor brake, is required to overcome static friction which becomes horsepower loss when running. At any given rate of feed-off, or weight carried, within reasonable range, friction-horsepower loss represents part of the energy which it is desired to dissipate and has little effect on the weight on the hook. The weight or speed controller reads the result from the hook load and applies the brake-controlling force to destroy part of the falling energy. The more friction-horsepower loss there is in the gear system, the less work there is for the brake to perform.

In control circuits, lost motion, leakage, false settings, or inaccurate instruments and controllers will cause errors. Fortunately, however, the control point is hand-set to a given visible indicator or gage point and, if the pickup device signal remains stable, and the instrument-brake combination does not hunt, the hook load will be maintained within close limits.

An error takes place in all weight-control systems which use either the dead line or torque from the live end, or drum end, as the point from which the weight signal is obtained. This is the weight of drilling cable added into the block lines as the kelly is lowered. On feeding down a 30-ft joint of drill pipe with eight lines, with the cable weighing 1 lb per ft, there are 240 ft or 240 lb of extra load supported by the cables when the joint is "down." This has the effect of adding this amount of weight to the bit at the lowest portion of the kelly. A device to adjust the instru-

[3] "The Relation of Nozzle Fluid Velocity to Rate of Penetration With Drag Type Rotary Bits," by J. D. Nolley, G. E. Cannon, and D. Ragland, *Oil and Gas Journal*, vol. 47, May 16, 1948, pp. 87, 88, 90, 93, 94, 97, 99, and 108.

ment control-point setting in proportion to the number of feet of line added would compensate for this error.

The only method of removing this error from the situation is to take the pickup signal from below the traveling block. This has been done, but adds complexities of having cables or hoses suspended in the derrick with means for keeping these connecting lines taut and preventing them from fouling the cables on the block.

When mud weights or characteristics or circulating rates change materially while drilling down slowly, a change in the self-balancing of the system will take place. In such cases, picking up of the drill pipe to check off bottom weight will determine the amount of change and resetting of bit weight necessary.

General Types of Control Systems Applicable to Drilling Rigs

Weight Control. Since weight on the bit is probably the chief determinant of drilling speed under given conditions, it is obvious that practically all systems of control have considered this problem.

Regardless of whether the weight on the bit is the primary control, or is governed by other physical effects, such as torque or feed speed, the weight must be controlled, or limited, because bits are built to withstand only so much weight, and still perform satisfactorily.

We have seen from previous sections of this paper that the hook-load signal is generally a sensitive and accurate means of determining bit load. Therefore it becomes obvious that a weight controller can be used and made to hold such weights as the driller feels desirable for the formation being drilled up to the limits prescribed by the bit manufacturer.

Under normal circumstances, in fairly homogeneous formations, a simple weight controller will suffice. In nature, rocks of very considerable hardness difference and drillability are layered often in inches or less, and a bit is generally designed to drill quite wide groupings of materials at constant weight. A considerably better performance is obtainable by the driller varying his rotary speed and weight on bit to known values to suit his experience. Weight control has the following features:

1 The surface signal is reliable and accurate and little affected by down-the-hole friction since the downward motion of the pipe is the "minor motion."

TABLE 4 GENERAL DESCRIPTION OF AVAILABLE AUTOMATIC CONTROLS

Make	Pickup signal	Controller and instrument-panel description	Recorder	Generator and panel
General Electric Company	*Weight:* null-balance current in motor armature *Speed:* tachometer generator voltage	Amplidyne control of motor current governs weight. Amplidyne control of field current controls speed. Separate speed and weight settings. Total load indicator. Bit-weight indicator. Supply volts. Control range setting, switches, 13-hp mill-type fan-cooled motor as brake, load-absorbing resistors. *General type:* Null-balance torque arm.	Bit weight and drilling rate with chronograph pen marking each foot drilled	Amplidyne motor generator set; 25-kw constant-speed generator or Thyrite. Constant-voltage generator, panel and controls
Dynamatic Corporation	*Weight:* null balance of excitation current in in field coils *Speed:* built-in tachometer generator	Electronic-type controller of field current from signal. Brake eddy-current type, includes mechanical reverse rotation of eddy-current ring by 5-hp motor at a speed greater than full-load slip speed. No-voltage trip. Overspeed trip. Emergency trip. *General type:* Null-balance torque arm.	Can be obtained to give same as General Electric	5-kw constant-voltage d-c generator with panel and controls, or a-c supply
OMSCO (Shell)	*Weight:* Pneumatic pressure from Hagan Thrustorq, also powers, and is in parallel with, Martin-Decker indicator gage *Speed:* Fluid pump against variable orifice. Remote orifice setting determines pressure at any speed. Pressure is speed signal to controller. True speed from tachometer generator	Weight and speed—Mason-Neilan 2500 pressure controllers. Brake, precision oil-immersed friction type. Tachometer generator and indicator. Speed controller is overspeed trip	Can be obtained to give same as General Electric	None. Requires about 3 cfm of air
Kinzbach Tool Company	*Weight:* Brake bands of drum on torque arms and supported on hydraulic cylinders. *Speed:* none.	Hand-control bleed rate from supporting hydraulic cylinders. Automatic pneumatic transfer and release from one brake band to the other. Air-operated main brake. Air puts brake "on" against spring pulley "off." Automatic control under development will be hydraulic bleed type	Special recorder being developed	Air compressors on separate unit

2 By providing suitable time delay for the instruments to operate against load changes, the load can be controlled accurately. It takes quite a considerable surface-feed distance to make any material load change on the bit, owing to the elasticity of the drill-pipe system.

3 The surface signal is extremely sensitive to bottom-hole weight changes under most conditions.

4 The penetration rate of the bit selects itself for the formation drilled. It is at the driller's control, secondarily, by changing the controlled weight, at his desire, to suit the formation.

Speed Control. A simple speed control on ordinary drilling operations is, in the author's opinion, an unsatisfactory single means of controlling the process. An ordinary hand or hydraulic brake, which can be set sensitively, is a form of speed control. Unless formations are quite homogeneous, and this is a rare state of affairs in well-drilling for more than a few feet at a time, a constant speed, slightly too fast, must "pile up weight" and, slightly too slow, must "drill off." As the speed-weight combination impressed by the controller crosses the speed-penetration-ability curve of the bit in that particular formation, stability of weight and speed can be reached, but with no relation to the load capacity of the bit. By being too light or too heavy, waste of bit life occurs on the one hand, or bit damage on the other.

On specialized operations, such as reaming, if formations are fairly uniform, speed control may be used to give good results, if down-the-hole stability is possible.

We have discussed the point that mean bottom-hole weight variations are very close to surface-weight variations. Surface variations of feed rate, however, due to the incalculable effects of elasticity in the string, have little relationship to down-the-hole speeds of progress, except that they are both in a downward direction.

Speed control has the following features:

1 The speed can be controlled only at the kelly. Both signal and control are done at the surface without relation to the drilling action.

2 The forward speed of the bit is self-determined and has no relation to the speed at the surface, except as interpreted by the driller on the weight indicator.

3 The weight on the bit has a value induced by the speed of feed and has no relationship to the bit manufacturers' recommendations, except as the driller changes the speed by hand.

Torque. A number of investigations and patents describe torque as a means of controlling drilling progress. This factor could be used as a control signal except for the fact that (*a*) torque changes take place when going from one formation to another, and (*b*) that a desirable drilling torque in one layer an inch thick may be too low for the next layer. This would require a torque increase. This can be applied to the bit only by lowering the drill pipe as fast as the torque change takes place. The next inch might require a reduction of torque and this would require lifting or stopping the pipe. This would be impractical and complicated, and it is believed better to let torque be self-determining within the reasonable strength capacity of the bit. The thought of torque as a control evidently stems from the old days when twist-offs were the bugbear of the driller. This is not the case nowadays and when trouble is experienced in this direction it is infrequently the fault of the bit. The drilling control cannot be expected to enable the user to forget fatigue, worn-out drill pipe or occasional "boll-weevil" stunts on the part of the crews.

If torque is used as any signal, it may be considered as a safety signal to disconnect the rotary in the event of too high values. This, or even a discussion of what excessive torque consists of for drill pipe, does not belong within the scope of this paper, and it is the author's opinion that any torque-control device is superfluous. A torque indicator, however, is a highly useful and desirable instrument for the driller's panel.

Combination Controls. While the foregoing discussion indicates that the weight control has by far the most desirable characteristics for general conditions, there are definitely occasions, such as occur when the bit suddenly penetrates a cavity, when the

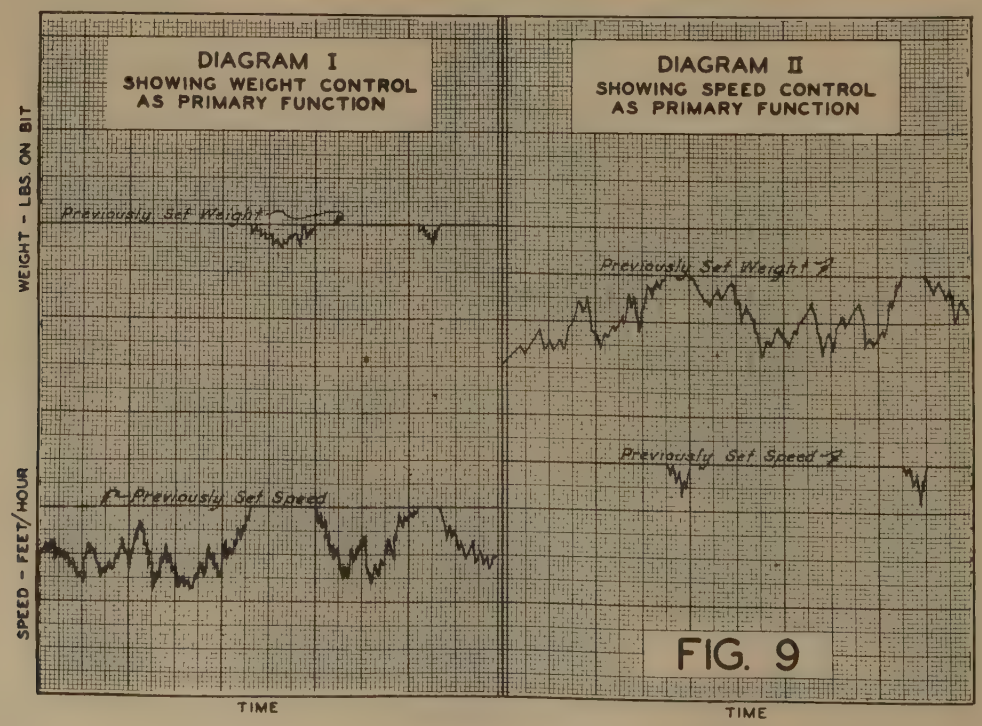


FIG. 9

simple weight controller would "run away." Also, when drilling very hard formations and a soft sand or very porous formation is entered by the bit, the same thing might happen. If this would not cause runaway, it might crowd the bit into the soft formation faster than the cuttings should be removed and jam the bit or "ball it up." In such cases, a speed "governor" with variable-speed-limit setting is highly desirable, in fact necessary, for safety's sake. For reaming or lightweight drilling, it may also prove desirable, and it is very useful for starting up the controller, since the weight-indicator reading, assumed to be the compensated self-balanced result of rotation, pump flow friction, etc., should be taken just before contacting bottom and with the bit being lowered at some controlled uniform speed close to the expected drilling speed.

Accordingly, and from experience obtained in the field, it appears that the bit-weight controller with variable-speed-limit control, also automatic, becomes the most useful controller type.

Overwind and External Power Sources. The braking device can be one of two distinct types. The first receives its rotative power from the wire line through the drum, and the brake merely destroys this energy. This is the simple overwind type and requires no external horsepower to operate the controller, other than the small amount needed for instrument operation. This involves $^1/_4$ hp at the most.

The second type is one in which the cable load is balanced against a torque produced by a device such as an electric motor. The balancing power comes from an outside source, or generator. While the first type is probably most desirable because of its ability to furnish its own power, the latter type generally can be made to give better sensitivity and speed of operation, and also can furnish motion in the reverse direction to lift the load slowly and reset the weight on the bit without using the main engines or drum clutches.

In determining which of the two devices will eventually be most acceptable, field experience as to requirements, cost, reliability, and simplicity, are the determining factors. The author feels that either type can be made sufficiently accurate and sensitive for the job, and that cost, reliability, and simplicity will be the main requirements.

Safety Devices

As with any other device of important duty, involving possible high monetary losses or danger to human life in case of failure, certain safety devices must be included in any good controller for a drilling rig. Since an acceptable control operates with the main brake handle completely up, the reason for safety devices is clearly seen.

Power Failure. In the event of failure of the power supply to the instruments or any part of the controller-supporting load a device must be installed to lock the controller brake or shaft. This would mean that, in case of spring-loaded brakes, the power must push the brakes "off" rather than "on."

The safety device must operate so that gradual leakage of, say, air from the system would apply the brake before the controller itself lost control of the load.

A low-voltage trip should also be included in electrical devices, which would not permit running unless proper voltage was available.

Overspeed Trip. In case of the operator turning a setting knob in the wrong direction and causing excessive speed to occur, an overspeed trip should be included. This device should operate the brake in a similar manner to the power-failure device.

Emergency Stop. A quick emergency stop should be available on the operator's panel, such as a valve or pushbutton, which will enable him to shut the brake off rapidly in case of emergency.

"Twist-Off" Indication. The weight controller has one useful characteristic, in that, in the event of a twist-off, less weight would be suspended on the hook and, regardless of whether the driller was present, this would have the same effect as too much weight on the bit, would apply the brake, and stop the feed. No further feeding would be possible unless the driller caused it to be so. This would eliminate much damage to the fish and prevent the troubles sometimes caused by "drilling by" with a twisted-off string. If the twist-off was caused by a fairly heavy section being broken off, the released spring effect in the pipe would raise the rotating portion several inches or more clear of the fish.

Application Methods

Transmission and Connection to Hoist. The controller transmission may be tied directly to the drum shaft or, preferably, may use any portion of the existing hoist reductions possible. In order to reduce static friction, which is undesirable in "starting" the brake, roller, or ball bearings are most desirable throughout. The use of chains is quite satisfactory, although gears are more compact.

Since the load factor at the highest loadings for a given unit is very small—probably below 5 per cent—normal transmission "wear ratings" can be considerably exceeded if the case or mountings are mechanically strong enough to resist occasional sudden brake application without distortion.

In order to allow the driller to throw the main engines into action in an emergency, the clutch to the control transmission should be connected with a nonreversible clutch. All clutch faces, if of jaw type, should be back-cut, since the whole string is suspended on these faces, and any vibrational throw-out would drop the pipe on bottom.

If a friction-connecting clutch is used, this should be control-interlocked with the main-engine-drive clutch so that the control transmission will be thrown out of action automatically before the engines are applied to the drum. Control of the transmission clutch must be at the driller's position.

Design of the transmission and brake setup requires torque, speed, and horsepower consideration.

Pickup or Signaling Devices. All weight and speed-control devices should be unaffected by, or compensated for, temperature changes within the operating range.

Table 1 shows a number of methods of obtaining suitable signals proportional to weight and speed. These signals may require amplification to give sufficient power to operate controllers. Differential devices may be used to amplify only the variations above and below set points.

Controllers. Weight and speed controllers should be practically unaffected by temperature over the operating range. They should be simple to operate and adjust by nontechnical personnel. Table 2 illustrates some types of controllers available.

Mixed types of controllers and pickups are very practical, and the simplest and most reliable signaling device for the job generally can be adapted to operate almost any type of controller the designer wishes to use. The controller in turn can operate several types of brakes.

Retarder or Brake. Table 3 gives a description of various types of brakes which can be used.

As pointed out previously, the transmission used is part of the energy-destroying system, and static friction of this system should be as low as possible. A very slow rate of creep is highly desirable so that static transmission friction is destroyed. This rate, however, should not exceed the equivalent of about 1 in. per hr feed rate of the brake, in order that it cannot affect the slowest drilling rates as, for instance, in chert when 2 to 3 in. per hr is quite common.

Such brakes as the eddy-current type require a mechanically reversed rotating field to overcome the natural slip of the brake

at low speeds. If hydraulic pumps are used with a slip rate greater than just prescribed, a make-up pump must be used to supply the slip fluid.

Any one of the forms of brake, except the electrical, which requires electrical control means, can be used with any type of suitable controller. The most suitable combination is one which uses the most reliable, rugged, simple and accurate device in each of the three main portions of the system.

Instruments for the System. In order to enable the operator to control the system and to locate troubles rapidly, certain minimum instrumentation is required, as follows:

1 Total weight indicator gage (existing standard gage may be used, if accurate). A weight-on-bit gage is desirable, reading 0—50,000 lb. A vernier scale is also desirable on the latter for accurate work, such as fishing.

2 A main power gage is necessary, such as voltmeter or pressure gage. This indicates value of voltage or pressure in the main power supply.

3 A rate-of-feed indicator showing feed rate in feet per hour.

4 Supply adjustment regulator and pressure gage for each main device used, such as weight and speed controllers, or motor valves used for other vital services.

5 Gages on controllers to show controlled variable, or signal, of speed and weight and output to motor valve or brake.

An additional instrument which is desirable is a drilling-rate recorder. This should include a record, on the same chart, of either the weight on the bit, or the total weight on the hook. It should also include a chronograph pen device which will mark the chart each foot drilled, and a switch on the driller's panel to allow the recorder to be started and stopped at the driller's will as the drill pipe is picked off or set on bottom.

Such a chart then becomes a penetration-rate chart for use of the engineers, and the inclusion of the bit-weight record shows whether a change in penetration rate was natural or caused by the driller having changed the weight setting on the bit.

This device can be included as a separate attachment and should be movable separately so that it can be used, or not. The recorder may be needed away from the driller's position and therefore requires a connecting cable, say, 75 ft long to give it flexibility of installation.

Economics

Cost Ranges. At the present time the cost setup for automatic controls is rather complex, since one of the main costs, up to 45 per cent of the total cost, is the "transmission cost." This varies considerably, depending upon the specific rig on which the system is to be installed. At this time, full descriptive literature on automatic controls is not available in bulletin type.

The figures in Table 5, assuming a maximum transmission cost, indicate a range of from $11,750 for the full-motor-reversing type down to $8500 for the mechanical-pneumatic type, including recording instruments.

The same control system for a medium rig, without the recorder would vary from $10,350 to $7250.

Profitability. At present, accurate data on profit obtainable by use of drilling controls is meager owing to the difficulty of obtaining directly comparable data.

The one directly comparative set of figures between two wells of the author's company in West Texas showed 43 per cent more footage per bit, 52 per cent more footage per day, and a saving of 113 bits.

No experience is yet available with the use of fishtail bits.

Trends. The increasing interest of many oil companies and manufacturers in automatic controls makes certain that, in due course, proper evaluation of the varied equipment offered and its effect on drilling costs will be made.

As drilling progress under similar conditions in the same fields with automatic and nonautomatic control can be compared, the situation will become clarified. More delicate but better bits for special purposes, such as diamond and special hard-faced bits, will, it is certain, accentuate the need for automatic control. It also is certain that the best service out of these highly designed bits cannot be obtained without precision and automatic controls to aid the driller.

Acknowledgments

The author wishes to acknowledge with thanks Shell Oil Company's permission to publish the data in this paper, and the assistance and co-operation of the General Electric Company, Gribbin and Baylor (Dynamatic Corporation), Oilfield Machine and Supply Company, and Kinzbach Tool Company, in furnishing data and illustrations.

References

"Principles of Industrial Process Control,"[4] by D. P. Eckman, John Wiley & Sons, Inc., New York, N. Y., 1945.

"Mechanical Measurements by Electrical Methods," by H. C. Roberts, Instruments Publishing Company, Inc., New York, N. Y., 1946.

"Servomechanism Fundamentals," by Henri Lauer, Robert Lesnick, and Leslie E. Matson, McGraw-Hill Book Company, Inc., New York, N. Y., 1947.

"Automatic Drilling Control," by M. E. True and R. R. Crookston, *Oil and Gas Journal*, vol. 46, Nov. 1, 1947, p. 70.

"Reaching Up to Go Down," *Oil Weekly*, vol. 118, July 16, 1945, p. 38.

[4] Terminology used in this report is in accordance with the ASME Industrial Instruments and Regulators Division Committee on Terminology; see pp. 224–230 in the reference.

TABLE 5 APPROXIMATE INSTALLED COST OF AUTOMATIC DRILLING CONTROLS

Make and type	Transmission: Class of rig	Transmission: Cost[a]	Controllers, brake, and driller's panel	Control	Recorder speed-weight combination	Power supply, generators and controls	Total[b]
General Electric Company	Heavy to Medium	$2500 to 1800	$6500 (Includes part transmission)	Automatic speed and weight	$750.00 (all sizes)	$2000	$11,750
OMSCO[c] (Shell) mechanical-pneumatic	Heavy to Medium	3500 to 2500 / 3000 to 2800	4250	Automatic speed and weight	750.00	Instrument air only from rig supply; 5 cfm max	8,500
Dynamatic Corporation	Heavy to Medium	2000 to 1800 / 1800 to 1500	6350 (Includes part transmission 100:1)	Automatic speed and weight	750.00	7½ kw d-c $1750	10,850
Kinzbach, mechanical-pneumatic	No transmission used		9000[d]	Nonautomatic[e]	750.00	1200	12,500 (automatic) 10,550 (nonautomatic)

[a] Variation depends on amount of hoist reduction usable in system.
[b] Heavy rig, maximum transmission cost.
[c] Being manufactured under license from Shell.
[d] Includes air-operated main brake, approximate value $2000.
[e] Automatic under development, probable extra cost—$2000.

"Shell's Rig Makes Successful Attack on West Texas Chert," *Oil and Gas Journal*, vol. 45, Oct., 12, 1946, p. 76.

"Tall Rig Designed for Drilling Deep Wells," by K. Marshall Fagin, *Petroleum Engineer*, vol. 17, July, 1945, p. 101.

Shell's Big Derrick and Its Record," by James C. Watson, vol. 7, April, 1946, p. 19.

Appendix

I Effect of Torsion in Drilling String on Length of String and Weight on Bottom

$$\text{Twist in degrees} = \frac{583\ TL}{N(D_1^4 - D_2^4)}$$

Where

T = torque, in-lb

L = length of string, in.

N = modulus of rigidity (shearing modulus) $^2/_5\ E$, or 12,000,000 psi

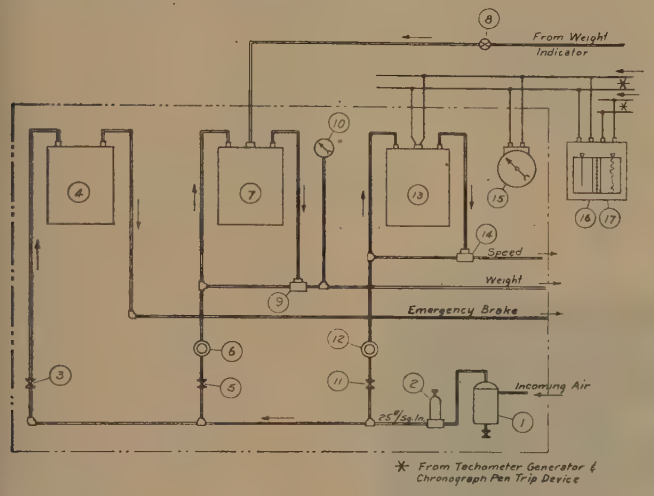
D_1 = outside diameter of pipe shaft, in.

D_2 = inside diameter of pipe shaft, in.

Example: Let us assume 100 hp carried by drill pipe at 100 rpm on 10,000 feet of drill pipe; $4^1/_2$-in. drill pipe, say, $4^1/_2$ in. OD, $3^3/_4$ in. ID.

Then

$$T = \frac{33{,}000 \times 100}{6.78 \times 100} \times 12 = 63{,}000 \text{ in-lb}$$



1 Air filter
2 Regulator
3 Valve
4 Emergency brake controller (regulator and pressure gage)
5 Valve
6 Regulator
7 Weight controller
8 Damper valve
9 3:1 air relay
10 Pressure gage
11 Valve
12 Regulator
13 Speed controller
14 1:1 air relay
15 Tachometer indicator (voltmeter)
16 and 17 Recording tachometer; chronograph pen

Fig. 10 Typical Control Panel, Mechanical—Pneumatic

1 Weight controller
2 Speed controller
3 Rate of penetration indicator

Fig. 11 "Shell" Type Automatic Control
(Control panel on left.)

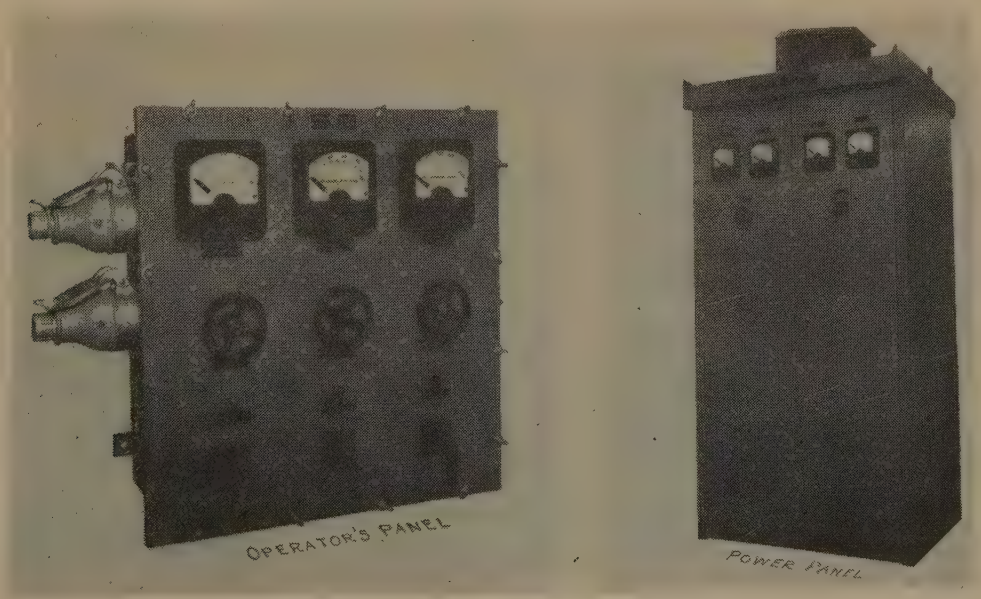


Fig. 12 Automatic Electric Bit-Weight Control
(General Electric Company.)

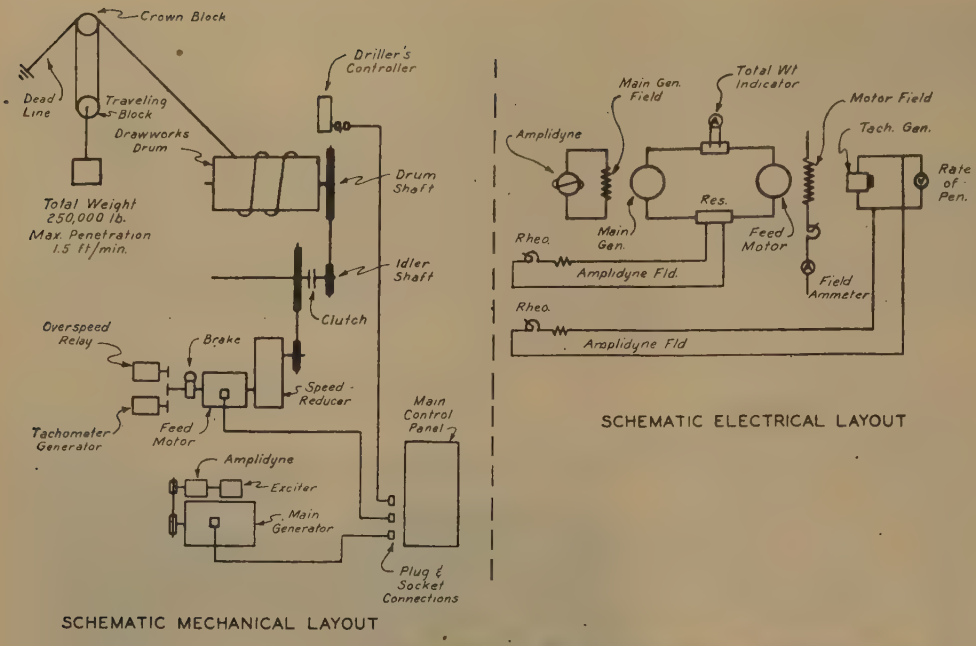
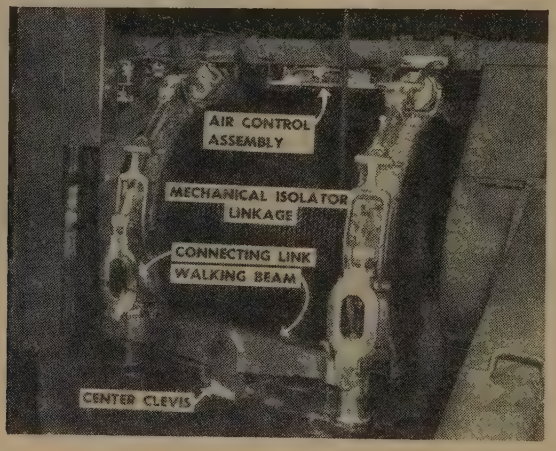
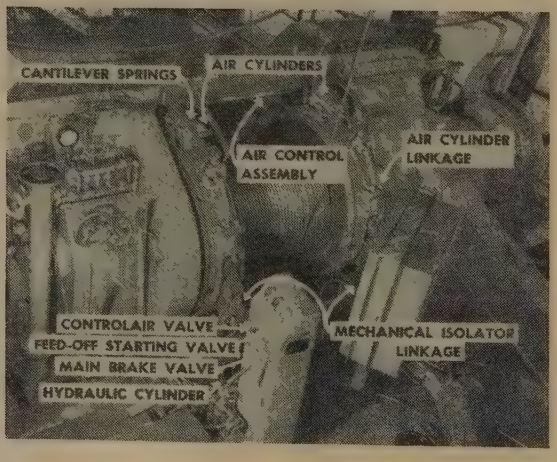


FIG. 13 AUTOMATIC ELECTRIC BIT-WEIGHT CONTROL
(General Electric Company.)



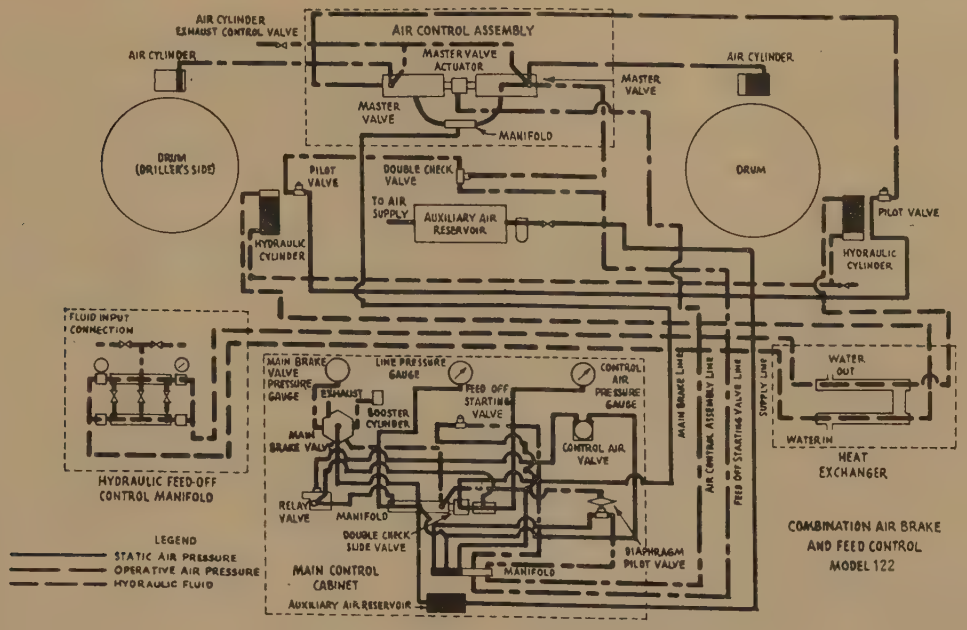
(a)

(b)

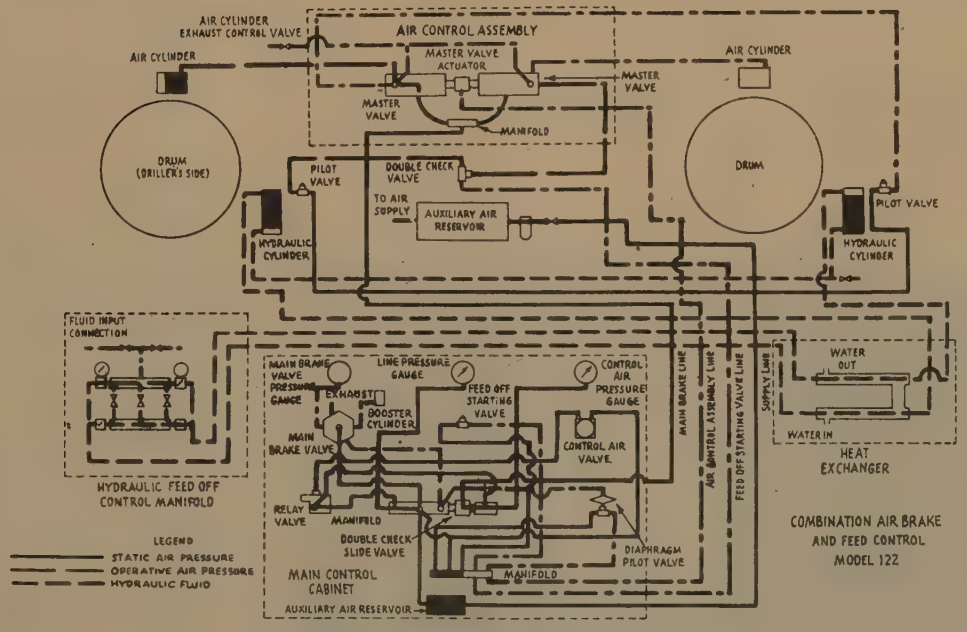


(c)

FIG. 14 WALKING BRAKE
(Kinzbach Tool Company, Inc.)



Application of Right-Hand Pilot Valve during Feed-Off Cycle



Application of Left-Hand Pilot Valve during Feed-Off Cycle

FIG. 15 FLOW SHEET WALKING BRAKE
(Kinzbach Tool Company, Inc.)

Therefore

$$\text{Twist} = \frac{583 \times 63{,}000 \times 10{,}000 \times 12}{12{,}000{,}000\ (4^1/_2{}^4 - 3^3/_4{}^4)} = 1718 \text{ deg}$$

or

$$\frac{1718}{360} = 4.77 \text{ revolutions}$$

Assuming 333 joints of pipe, this is 1718/333 deg per joint, or 5.15 deg.

Distance subtended on outside diameter 5.15 deg torsion =

$$4^1/_2 \times \pi \times \frac{5.15}{360} = 0.202 \text{ in.} \quad \text{On a 30-foot joint}$$

$$\text{Longitudinal angle of twist} = \sin^{-1} \frac{0.202}{360} = 0.000561 \text{ or approx } 0° 1.9'$$

$$\text{Shortening of joint by torsion} = 360 - \sqrt{360^2 - 0.202^2} = 360 - 359.9999 \text{ or less than } {}^1/_{10{,}000} \text{ in.}$$

This is less than 0.03 in. in 10,000 ft, or negligible.

Therefore, since length of string is not affected, weight on bottom is, for all practical purposes, unaffected by severe torsional variations.

II Wall Friction and Hook Load

In a well, friction forces F exist between well wall and drill pipe.

Vertical Movement. Friction load ΣF becomes measurable. Its value is obtained by finding weight reading W_1, moving when downward at a slow rate, and weight reading W_2, moving "up" at the same speed

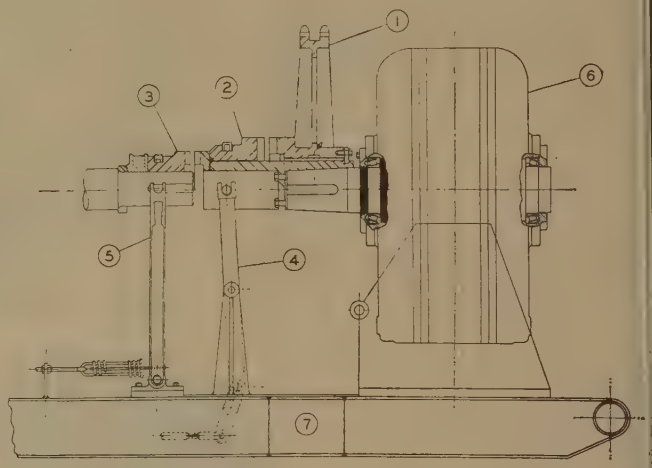
$$\text{Friction load} = \frac{W_2 - W_1}{2} \text{ lb}$$

$$\text{Friction hp} = \frac{W_2 - W_1}{2} \times \frac{V}{550}$$

where V = velocity of movement, fps.

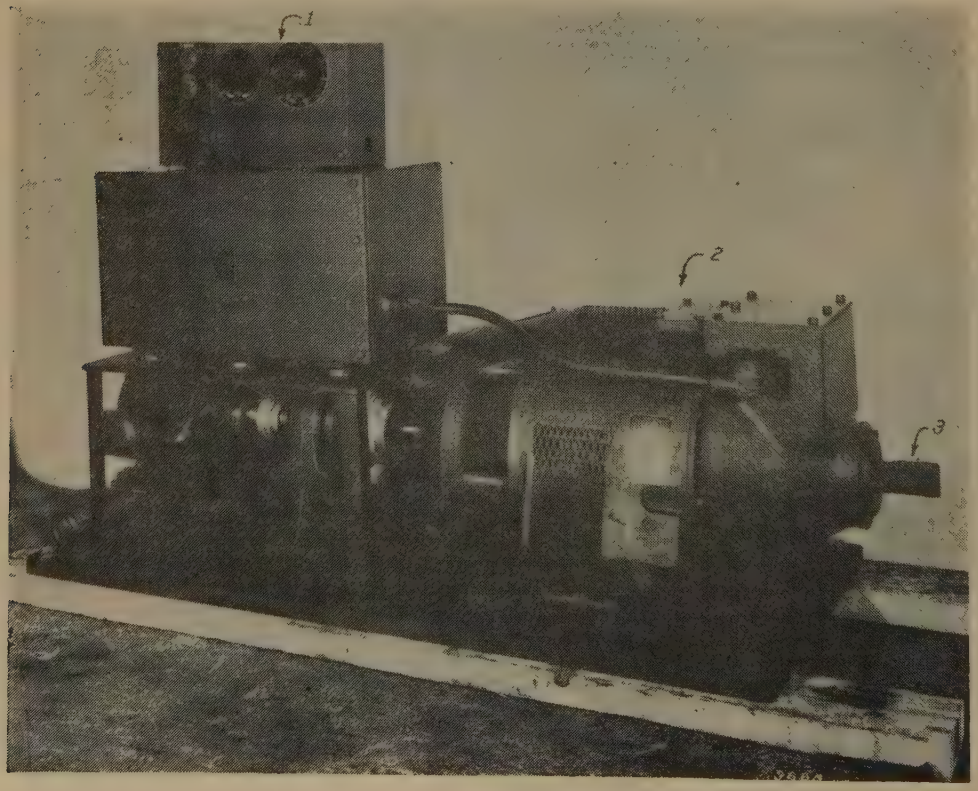
Torsional Movement. Friction load F, can be converted into friction horsepower loss by applying a torsional force T at a rate N

$$\text{Friction horsepower} = \frac{2\pi NT}{33{,}000}$$



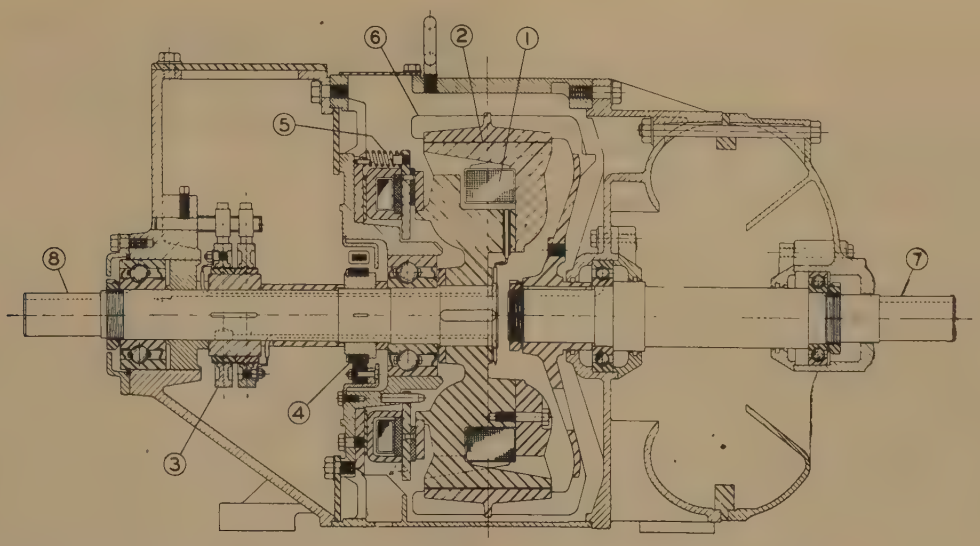
1 Drive to automatic control transmission
2 Automatic control disconnect clutch
3 Hydromatic brake disconnect clutch
4 Shifter automatic control clutch
5 Shifter hydromatic brake clutch
6 Standard hydromatic brake assembly
7 Drawworks' skid extension for automatic-control drive

Fig. 16 Typical Drive Drumshaft to Automatic-Control Transmission



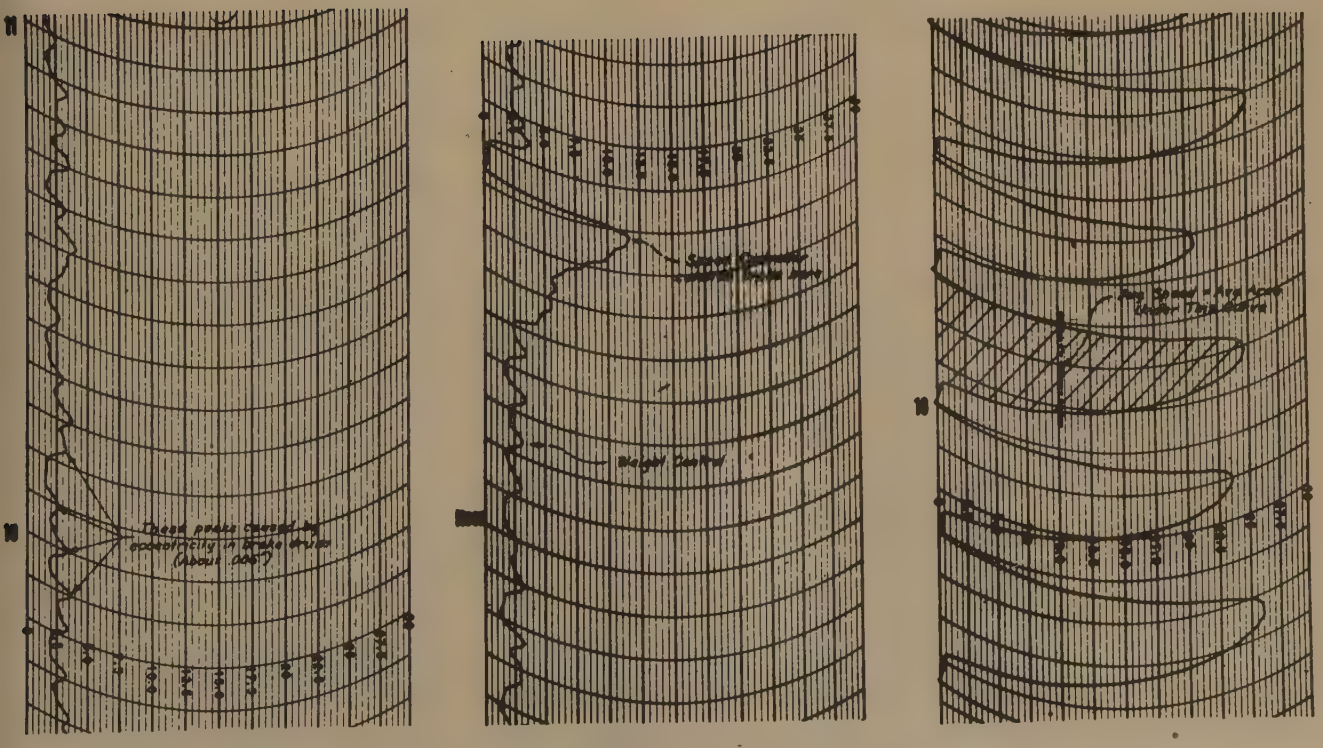
1 Driller's control panel
2 Brake unit
3 Connection transmission

Fig. 17 Automatic Drilling Control
(Dynamatic Corporation.)



1 Main excitation coils
2 Air gap
3 Slip rings to main excitation coils
4 Tachometer generator
5 Power-failure emergency-stop brake—applied if power fails
6 Cooling fins
7 Antislip drive, driven in reverse to overcome full-load slip
8 Hoist-drum drive

FIG. 18 DYNAMATIC AUTOMATIC DRILLING-CONTROL BRAKE UNIT
(Gribbin & Baylor.)



Brake on 100 % Weight Control
Chart Speed: 12" per Minute

Weight Control with Speed Control Limit
Chart Speed: 12" per Minute

100 % Speed Control
Chart Speed: 12" per Minute

FIG. 19 TYPICAL RECORDER CHARTS; AUTOMATIC DRILLING CONTROL

where

N = rpm
T = torque, ft-lb

In torsional movement, the lubrication of the rubbing surface by drilling fluid is good, and the work done is proportional to distance traveled per unit of time, or the velocity of rubbing. All friction is thus converted into horsepower loss. This is the major motion of the drill pipe.

In vertical movement, while rotating the pipe, vertical friction load, which has been redetermined by rotation, must be converted into horsepower loss. This is the minor motion of the drill pipe.

Frictional horsepower is proportional to rubbing velocity, since the frictional resistance is the same against vertical and horizontal motion.

Example: $4^1/_2$-in. drill pipe with $5^3/_4$-in-OD tool joints, running at 150 rpm, and drilling at a rate of 120 fph.

Peripheral velocity of rubbing surface $= \dfrac{5^3/_4 \times \pi \times 150}{12 \times 60} =$ 3.75 fps.

Vertical velocity $= \dfrac{120}{3600}$ fps $= {}^1/_{30}$ fps. Therefore peripheral velocity = 112.5 times vertical velocity, even at the abnormally high drilling rate of 120 fph.

Vector Diagram of Forces: In Fig. 20, θ = pitch angle of a thread cut by a point on the outside diameter of the $4^1/_2$-in. drill pipe; in this case, 150/24 threads per inch or 6.25 threads per inch.

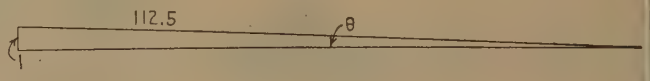


FIG. 20

This situation is relatively unchanged on bottom, and as long as peripheral velocity is of the order of 50 or more times that of vertical velocity, vertical friction is removed from the situation, for all practical purposes.

Close observation in the field shows the foregoing to be true, in that, when the pipe is suddenly lowered, while rotating, and then stopped, the vertical friction can be seen to "run down to bottom," i.e., the weight indicator shows a sudden drop which gradually, over a few seconds, rises to its original rotating value. When lowered by the controller at a steady rate while rotating at normal speed, however, the indicator shows no loss of weight.

Pressure Surges and Vibration in Reciprocating-Pump Piping

By J. W. SQUIRE,[1] TULSA, OKLA.

In the early days of the petroleum industry in America, the long pipe lines through which oil was pumped gave constant trouble from bursts caused by "shock" pressures. The author traces pump and other developments through the years which were intended to reduce the surge problem. Practically all solutions were based upon the use of air chambers, equalizer lines, or otherwise designing piping to prevent these surges from becoming dangerous. With the pumps in use today, pressure swings of 50 to 60 per cent of the mean pressure are experienced. Steps taken to so arrange the flow in the pump that these pressure surges are held within a reasonable percentage of the theoretical value are outlined in the paper. These consist mainly of changes in the pump-valve system, which the author describes at some length, supported by mathematical and test data.

History and Introduction

DURING the growth of the pipe-line industry, the problem of surges in the lines because of fluctuations of pressure at and near the prime movers has plagued the operator. In 1903 Professor Goodman[2] recognized some of these problems in the following statement:

"In the early days of the petroleum industry in America, the long lines through which the oil was pumped gave constant trouble owing to frequent bursts, although the pipes had a large margin of safety over any static pressures to which they were likely to be subjected. When precautions are not taken to prevent sudden changes of velocity in the pipes and passages of fast-running pumps, very serious 'shock' pressures may be set up even with short lengths of pipe."

Most of the early crude-oil pumps used were direct-acting, operated by steam. The steam cylinder cushioned the shock somewhat and few breakages were experienced. The next step was using a steam crank with a light flywheel directly connected to the pump. The steam end still cushioned the shock effects of the pressure surges. The widespread use of the slow-speed Diesel engine using heavier flywheels and a gear-reduction system to the pump in the early 1920's served notice to the pipe-line industry that further consideration of the surge problem was necessary. What had been manifesting itself in "short stroking" in the steam-operated units now began to appear as numerous and widespread breakages of the lines. When the lines and line fittings were "beefed-up," the pumps broke. When the pumps were strengthened, the lines broke. It was not until about 1924, therefore, that the seriousness of the problem became apparent. There were then some 40,000 miles of pipe lines in the United States carrying crude oil.

[1] Stanolind Pipe Line Company.

[2] "Hydraulic Experiments in a Plunger Pump," by John Goodman, Proceedings of The Institution of Mechanical Engineers, London, England, 1903, pp. 123 and 172.

Contributed by the Petroleum Division and presented at the Petroleum Mechanical Engineering Conference, Amarillo, Tex., October 3–6, 1948, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—PET-9.

At this time Mr. W. D. Pomeroy, vice-president and general manager of the Goulds Manufacturing Company, Prof. H. Diederichs, director of the Sibley School of Mechanical Engineering, Cornell University, and others instituted a study of the causes of these breakages. Their research was concerned primarily with an attempt to adapt the use of air and vacuum chambers or alleviators on the suction and discharge sides of the pumps to cushion the shock of the pressure waves set up by the pump in the piping systems in use. In so doing, a study was made of the effects of the piping-system design on these pressure waves. This work was carried on both at the Goulds factory in Seneca Falls, N. Y., where it was feasible to study test units under controlled conditions, and in the field under actual operating conditions at several pipe-line stations. A preliminary paper outlining the results was presented in 1926 by N. B. Delavan.[3] Their main results and specific conclusions were given in a paper presented in 1929,[4] and are summarized as follows:

1 If the piping design is such that a reflecting point, such as a sudden enlargement or contraction of the line, carries back pressure waves which are in synchronism with those of later waves of pressure produced by the pump plungers, very high pressures are developed and failures of the line or pump parts may result. This phenomenon has long been known as water hammer and many studies have been made of it. Field conditions were found where pressure swings of 310 to 1030 psi on a 675-psi line were typically present. When these conditions exist to such an extent that the liquid column breaks, they are even more serious.

2 The pressure impulses travel at the speed of the sound in the liquid, being modified by the gravity of the liquid and the size of the pipe

$$C = \sqrt{\frac{gJ}{m}}$$

where

C = velocity of pressure wave in system, fps
g = 32.2 fps per sec
m = density of liquid, pcf
J = virtual modulus of elasticity of liquid and pipe (see Appendix)

Average figures of C for oil are around 4000 fps in 6 and 8-in. pipe as derived in this paper

3 Short lengths of piping between the pumps and manifold are desirable and cause less breakage than long lines, particularly when more than one unit discharges into a common manifold. The conditions vary considerably where long lines between pumps and manifold exist; in fact, no two are the same. If there are no critical points of reflection, the primary impulses tend to die out because of friction loss in the pipe and internally in the liquid. Any actual installation, however, is seldom without

[3] "An Investigation Into the Cause of Breakages in Crude Oil Pipe Line Transportation Systems," by N. B. Delavan, Trans. AIME, vol. 73a, 1925, pp. 498–520.

[4] "The Occurrence and Elimination of Surge or Oscillating Pressures in Discharge Lines From Reciprocating Pumps," by H. Diederichs and W. D. Pomeroy, Trans, ASME, vol. 51, 1929, Paper No. PET—51-2.

discontinuities in the piping, which multiply the chances of a resonance condition in the piping system, particularly when it is considered that all odd harmonics must be taken into account.

4 The use of air chambers near the pump, when intelligently applied, prevent dangerous surges.

The foregoing points represent those of the authors previously mentioned and are amply and completely covered in their paper. It is not the intent of the author to do more than summarize briefly their conclusions. Much information and valuable data are included in this work in particular for speeds and design of pumps up to that time. The latest units then consisted of duplex or triplex double-acting pumps powered by Diesels and using a positive gear-reduction drive.

Surge chambers were used with some effectiveness; but the ever-present possibility of an explosion, the extra cost of installation, and the cost and inconvenience of maintenance were disadvantages to be considered. The construction of additional larger units at all main-line stations with all pumps working into a common manifold resulted in numerous piping failures.

In order to attack this problem further, J. B. Harshman of the author's company made an investigation for the purpose of devising a simpler and more economical means of reducing pressure-resonance conditions in multiunit installations. The results of his work were summarized in a paper in 1939.[5]

The method consisted of connecting an equalizer line to each discharge line of the pumps. This was placed a short distance from the pump-discharge headers just outside the building. Practical tests in the field showed a 1350-psi pressure swing on a single unit in a four-unit station with three units discharging at 725 psi line pressure and connected without the equalizer line. The insertion of the equalizer line under otherwise the same operating conditions reduced this pressure swing to 120 psi, a 91 per cent reduction. Several other tests gave similar results. On two to four (or more) unit stations, this means for reducing pressure surges or fluctuations has been used effectively in numerous installations since that time.

Recently, however, several other factors have appeared which bring up new problems. One pipe-line company now has a number of 5- and 7-cylinder vertical and twin-quintuplex horizontal pumps. These range in output from approximately 15,000 bpd to 100,000 bpd, in speed from 175 rpm to 400 rpm, in stroke from 3 in. to 8 in., and in output pressure from 900 psi to 1500 psi. The trend then is for higher speeds, greater capacities, higher pressures, and lighter weight in the pumping units.

The first of these, put into use in 1944, gave some indication of the difficulties which were to be experienced. However, there were no piping-breakage difficulties on the first few installations; and the somewhat excessive noise was attributed to the speed at which the pumps were running. Some changes were made in the plungers and valve springs to reduce this noise. With the installation of larger and faster pumps of the same design in 1947, piping breakage began to appear. It was then that a series of tests was instituted with equipment which will be outlined later.

All those who previously had carried out work on crude-oil pumps, recommended measures for alleviating the pressure surges by the use of air chambers or equalizer lines, or otherwise designing the piping to prevent these surges from becoming dangerous. The nature of the surges produced in the pump cycle for the slower duplex and triplex pumps was that of large and slow variations. Now, however, the theoretical pressure variations generated by the action of the pump were small and of high frequency. Reference to Table 1 and Figs. 1 to 5, inclusive, demonstrates this clearly. But instead of experiencing these small pressure variations in the field, measurements were made of

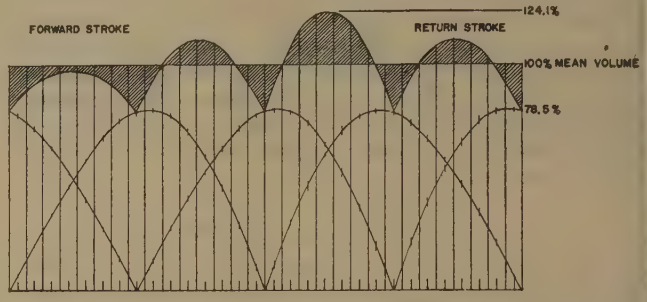


FIG. 1 DISCHARGE VOLUME VARIATION OF DUPLEX DOUBLE-ACTING PUMP

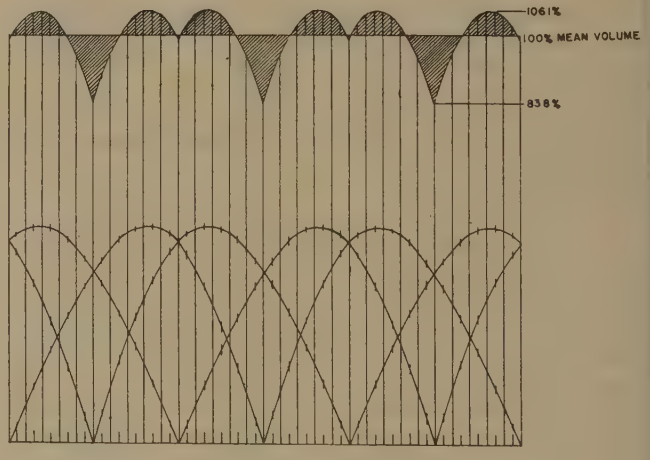


FIG. 2 DISCHARGE VOLUME VARIATION OF TRIPLEX DOUBLE-ACTING PUMP

[5] "A Practical Method of Arresting Pressure Surges in Discharge Piping of Reciprocating Pumps," by J. B. Harshman, *Oil and Gas Journal*, April 27, 1939.

TABLE 1 FUNDAMENTAL VOLUMETRIC VARIATION OF TYPICAL PUMPS

Type pump	Rpm	Per cent volumetric variation: Above mean	Below mean	Total	Cycles per rev	Cycles per sec	Reference
Horizontal double-acting duplex	36	24.1	21.5	45.6	4	2.4	Fig. 1
Horizontal double-acting duplex	58	24.1	21.5	45.6	4	3.9	Fig. 1
Horizontal double-acting-triplex	36	6.1	16.8	22.9	6	3.6	Fig. 2
Horizontal double-acting triplex	73	6.1	16.8	22.9	6	7.3	Fig. 2
Vertical quintuplex	175	1.8	5.2	7.0	10	29	Fig. 3
Vertical quintuplex	277	1.8	5.2	7.0	10	46	Fig. 3
Horizontal twin quintuplex	425	1.4	4.2	5.6	10	71	Fig. 4
Vertical septuplex	225	1.2	2.8	4.0	14	52.5	Fig. 5
Vertical septuplex	256	1.2	2.8	4.0	14	60	Fig. 5
Vertical single-acting triplex	229	6.1	16.9	23.0	6	22.9	...
Single-acting quad	...	11.0	21.5	32.5	..	...	...
Single-acting sextuplex	...	4.8	9.2	14.0	..	...	...

total pressure swings on the order of 50 to 60 per cent of the mean pressure—a factor of 8 or 9 times the expected value. Our investigations indicated that air chambers or equalizer lines could not be used effectively at these frequencies (see Appendix).

The method of attacking the problem, then, involves primarily the concept of so arranging the flow in the pump that these pressure surges are held within a reasonable percentage of the theoretical value. Then too, the piping should be so designed that no part of it will approach a resonance with these surges initiated in the fluid cylinders of the pump.

This means that instead of attacking the problem by alleviating means, we are attempting to stop any dangerous surges at their source. Fig. 6 shows a general layout of the pump and piping system where the conditions just mentioned were experienced. Fig. 7 shows one record of the pressure surges taken before any changes were made in the pump. This has superimposed the volumetric curve of Fig. 3 converted directly in terms of pressure.

After changing the valve-system design and introducing other factors, recorded results were of such low value at this point

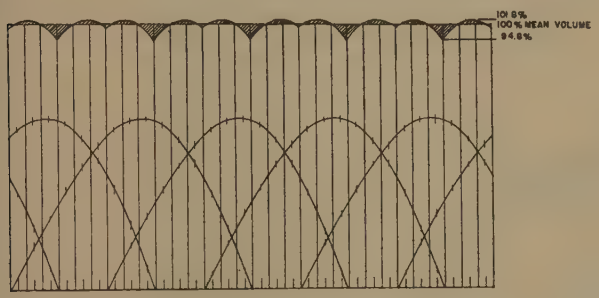


FIG. 3 DISCHARGE VOLUME VARIATION OF QUINTUPLEX SINGLE-ACTING PUMP

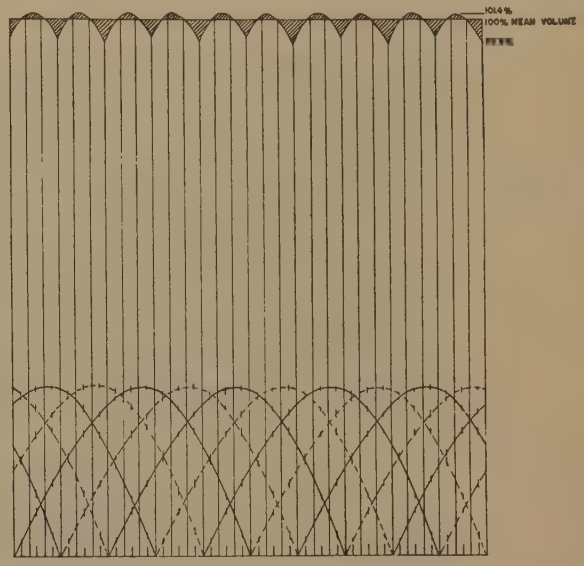


FIG. 4 DISCHARGE VOLUME VARIATION OF TWIN QUINTUPLEX PUMP

FIG. 5 DISCHARGE VOLUME VARIATION OF SEPTUPLEX SINGLE-ACTING PUMP

that the character of the curve was not believed recorded faithfully by our equipment. This is illustrated in Fig. 8, again with the theoretical volumetric variation shown. The maximum variation of 39 lb on the 3000-psi gage used, amounted to a little over 1 per cent of its range.

What produced this marked improvement in operation? No changes were made in the piping. The changes in the pump are enumerated as follows:

1 Differently designed valves and valve seats were used in three of the five cylinders. These were such that discontinuities in the velocity of flow through the seat and across the rim of the valve were reduced considerably. At the same time, the lift of

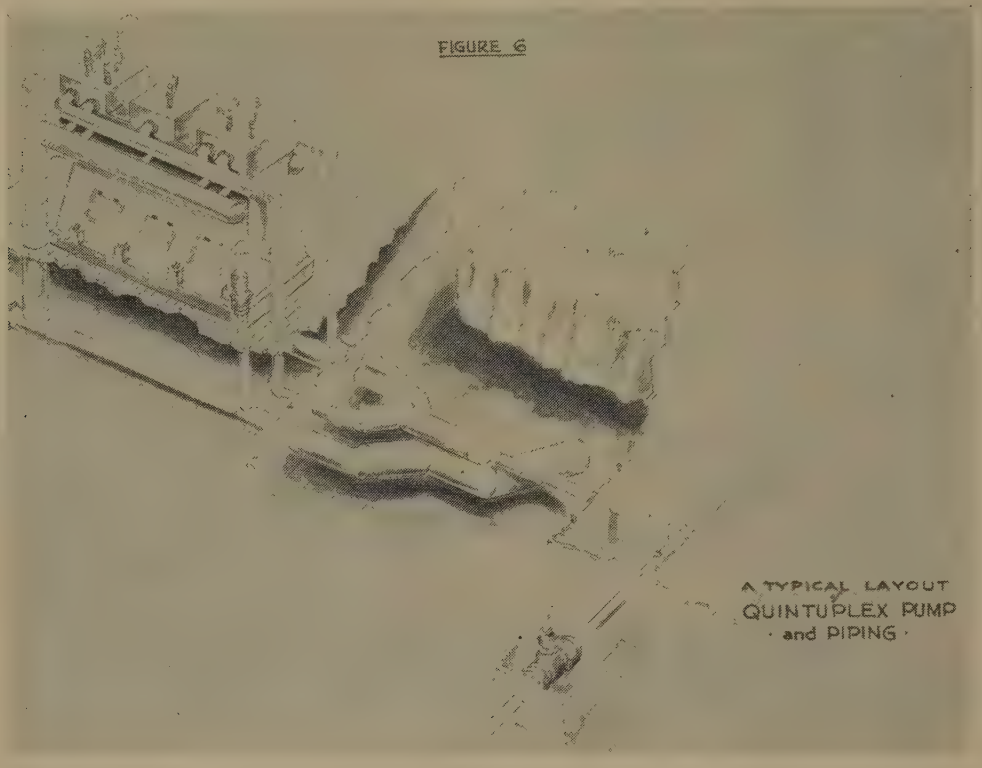


FIG. 6 TYPICAL LAYOUT OF QUINTUPLEX PUMP AND PIPING

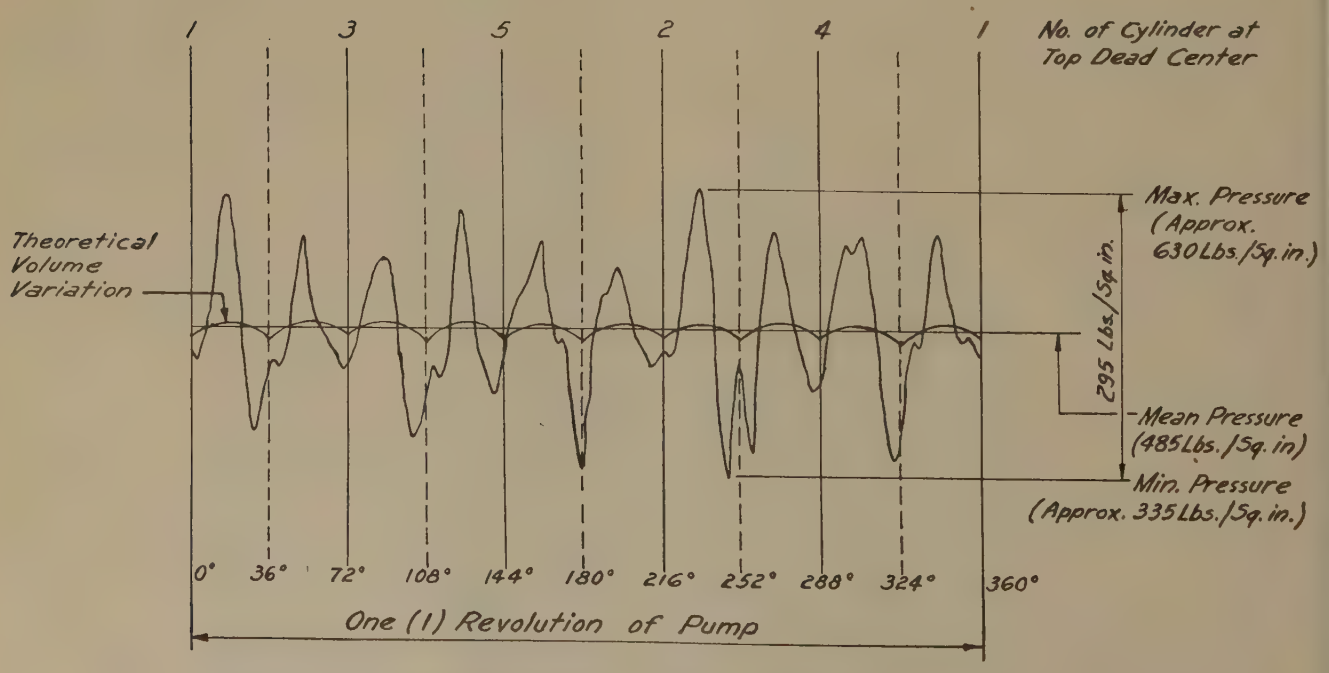


FIG. 7 DISCHARGE PRESSURE VARIATION APPROXIMATELY 8 FT FROM DISCHARGE MANIFOLD OF VERTICAL QUINTUPLEX PUMP

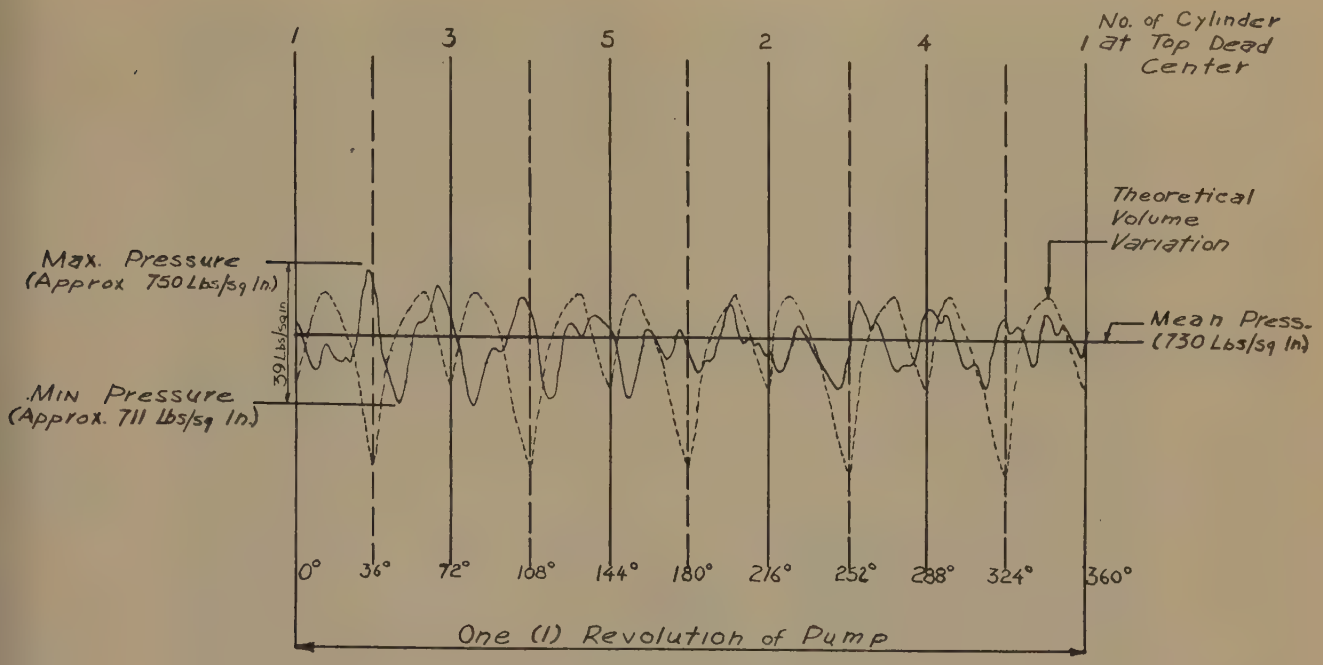


FIG. 8 DISCHARGE PRESSURE VARIATION APPROXIMATELY 8 FT FROM DISCHARGE MANIFOLD OF VERTICAL QUINTUPLEX PUMP

the valve was restricted to 1/8 in. in all cylinders. (No lift restriction had previously been imposed.)

2 The new valve systems had a higher natural frequency by a factor of approximately 3.

3 Diameter of the plungers was changed from 7 3/4 to 8 in.

It is believed that the reduction of discontinuties in the velocity of flow reduced the forces likely to excite the valve system into resonance. This, coupled with the higher natural frequency of the valve system, was credited with being the main factor resulting in greatly improved operation of the pump. The valve-lift reduction may also have been a contributing factor.

SCOPE OF THE PROBLEM

The Pump. This leads us to a consideration of the following general factors in the design of high-speed multicylinder pumps for crude-oil service:

1 The pump parts must be built to withstand greater strains; such as, centrifugal forces in the crankshaft and greater bearing pressures, due to increased line pressures.

2 The moving parts must be built to closer tolerances, and a good dynamic balance must be attained.

3 The design of the valves and valve seats must be such that efficiency of the movement of oil is at a maximum. Theoretically, sharp velocity changes in the flow are apt to cause valve flutter. Also, theoretically it should not be necessary to restrict valve lift beyond the maximum to produce equivalent port area in a properly designed valve. However, practical tests by several pump companies have resulted in imposing a limitation on the valve lift. This lift must be kept reasonably small to offset the inertia forces which prevent proper closure, thereby allowing quick action of the valve. Closure at the proper time means increased volumetric efficiency. Small lift also means less noise and wear on the valve plate and seat. The amount of lift on a properly designed valve should be in direct proportion to the speed of the plunger. Such a valve would be wide open when the crank is 90 deg to the plunger and within 0.010 in. of being closed at the end of the stroke with maximum pump speed.

4 Therefore the valve system should be light in weight and designed so that its natural frequency is not close to the frequency of the hydraulic forces acting upon it. This obviates fluttering.

5 The spring system on the suction and discharge valves must be designed to handle a wide variety of oils with minimum adjustment. Complex batching schedules frequently result in wide variations in back pressures and sometimes in discharge pressures. The compressed force of the springs must account for these factors as well as the weight of the valve stem and spring for optimum operating conditions. The ideal valve would have neither weight nor mass and would depend entirely upon spring pressure for closure. Since this is impossible, the mass should be kept to a minimum with ample spring pressure. In other words, the G factor should be kept to a minimum.

The designed factors 3, 4, and 5 involve many considerations on which little known data are available. A possible answer in design of future units would be the use of a system to actuate the valves by a mechanically positive means. This would obviate consideration of the many varying factors which must now be taken into account in the design of a valve system hydraulically operated.

Considerable field experimentation has been necessary to arrive at satisfactory operation of the spring-system design on units now in use. These tests are by no means complete.

Piping. Principles which are used in designing the pump-discharge and suction piping may be summarized as follows:

1 Streamlining of flow wherever possible. Examples of this may be seen by reference to Figs. 6, 9, and 10. In the piping in Fig. 6 it will be noted that an attempt has been made to equalize the streams on both the suction and discharge sides of the pump to avoid unbalance in its operation. Fig. 9 shows the cast discharge-header construction now in use on the pump in Fig. 6 and on several others on our system. Experience in the field has dictated the policy of straightening out the flow, particularly in the discharge lines. Fig. 10 illustrates an example of bringing together the streams from a double-end-discharge vertical septuplex pump. This pump has been operating satisfactorily at

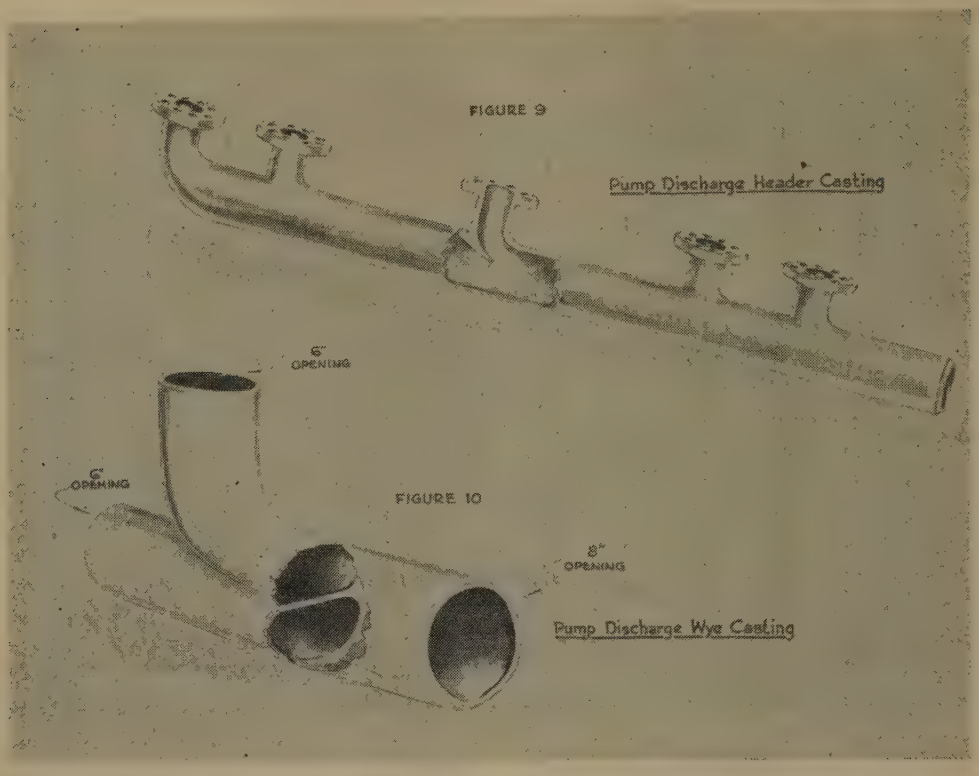


FIG. 9 (*above*) PUMP-DISCHARGE HEADER CASTING
FIG. 10 (*below*) PUMP-DISCHARGE WYE CASTING

1200 psi discharge pressure for several months with no surge chambers. Piping vibration has been negligible.

On the other hand, avoiding the use of long-radius manufactured ells ($1^1/_2$ times nominal pipe diameter) and manufactured tees is considered impractical. The chances of reflections in synchronism with the pressure surges initiated by the pump, although possible, are believed to be negligible for two reasons; (1) these fittings are very imperfect reflectors; (2) there are usually more than two points in the line between the pump and manifold where the use of one or both is necessary. These facts taken together probably mean that the resultant pressure waves are damped out to such an extent that they are not of destructive nature. Our measurements in the field have indicated this to be true, although we have been unable to calculate what is happening in any particular case.

2 Avoidance of suspended masses of piping or fittings. Small surges initiated by the pump are apt to approach the fundamental natural frequency or one of the harmonics of some section of the piping with destructive results.

3 Avoidance of irregularities in cross section, such as, holes, sharp shoulders, or notches. We try to use a minimum number of saddle welds—never size on size with pressures above 500 psi. Where economic or delivery reasons dictate their use rather than the purchase of a factory tee (e.g., 6 on 16 in.), they are usually rib-reinforced at the crotches. All discharge gage connections are rated at 2000 psi for strength.

It is also possible for mechanical vibrations to be transmitted to the piping from the pump or engine, causing dynamic stresses with which some part of the piping system is in resonance. When this occurs, breakage results. Several studies have been made of this phenomenon in the natural-gas industry. This is a related problem to the one covered in this paper, but we have not as yet experienced any difficulties in piping breakage attributable to this cause.

Fundamentally, then, the discharge-piping system is subjected to pulsations of much higher frequency and less magnitude than any heretofore experienced for crude-oil pumping. Referring to Table 1, it will be noted that the pulsation variations produced by the pump vary between 2.4 and 7.3 cps for horizontal duplex and triplex double-acting pumps whereas in the multiplex pumps this figure varies between 29 and 71 cps. The theoretical volumetric variation per revolution on the other hand (and it may be assumed the pressure variation follows this in character on the pump discharge) varies from 45.6 to 22.9 per cent for the duplex and triplex pumps, respectively. The comparative variation for our multicylinder pumps in operation varies from 4.0 to 7 per cent. This means:

1 That the number of flexures of the piping system in a properly operating multiplex pump occur at a rate of from 4 to 30 times as fast as those of the 2 and 3-cylinder pumps with an average factor of about 10.

2 The amount of the pressure surges are at the same time much smaller in the multiplex pumps than in the duplex and triplex pumps. This again is for a pump operating ideally.

3 The theoretical conditions of pressure are seldom fulfilled in actual practice. It is felt necessary, therefore, to design the piping system to handle unforeseen pressure surges at the present time.

When working into discharge pressures in the range of 1400 to 2000 psi and handling crudes with a wide variation of viscosity,[6] the safety factor in the discharge-piping system is usually smaller, e.g., for an 8-in. discharge line designed for 1400 psi operation using an S factor of 18,000 psi,[7] the allowable pressure is 1640 psi using extra-heavy grade "B" seamless pipe. This means, then, that the allowable working pressure should not vary over

[6] 40 SSU to 10,000 SSU for certain cases in Wyoming.
[7] 85 per cent of the S factor in the ASA code for oil-piping systems within refinery limits.

17 per cent above the mean. On the other hand, in an 800-psi discharge line using 8-in. pipe, extra-heavy grade B seamless would again be used, giving an allowable percentage of variation of 105 per cent above mean operating pressure.

Measuring Equipment. Fig. 11 shows a schematic layout of a system satisfactory for pressure and vibration measurements in a high-speed crude-oil pump-and-piping system.

In arriving at specifications for this equipment, several factors must be kept in mind, as follows:

1 We must have a system capable of following the pressure variations faithfully. We have set our standards arbitrarily at 0 to 1000 cps.

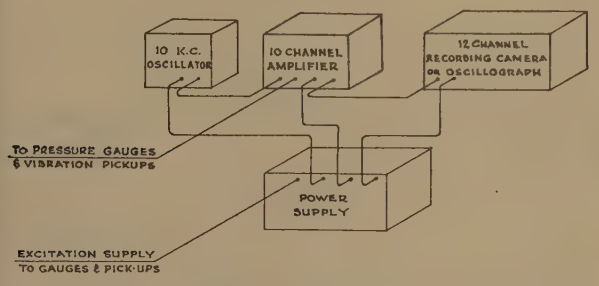


FIG. 11 . CABLING DIAGRAM OF VIBRATION AND PRESSURE-MEASURING EQUIPMENT

2 The gage for transmitting the pressure impulses also should be as insensitive as possible to mechanical vibrations. This is one of the main difficulties which we have experienced in use of pressure gages with an electrical output. As yet, we have found no suitable gage fulfilling our frequency requirements and vibration-insensitivity requirements.

3 It is believed highly desirable to measure the direction and amount of mechanical vibration encountered in the piping system simultaneously with the measurement of pressure. This means using a pickup for each of three directions of motion at the measuring point. The frequency characteristics in this case have been set (again arbitrarily) at 4 to 500 cps.

4 The recording equipment should be capable of handling data from a minimum of two vibration points and four pressure points simultaneously on the same time scale. This is a total of ten channels; six for the two points of mechanical displacement (one for each direction of motion at each measuring point), and four for the pressure. Data should be directly readable from the record in terms of inches displacement and pounds per square inch pressure variation. A recording oscillograph or camera such as is used in seismograph work is suitable.

This type system also allows us to observe conditions quantitatively for several consecutive pump revolutions allowing such factors as hunting of the engine or beats in the data to be observed and evaluated.

The requirements of the system briefly described here have been arrived at after a series of tests conducted on one of our single-unit pumping stations in Missouri. It is believed that by proper design all of the time-consuming factors which we experienced in taking and interpreting the data can be minimized, allowing a quick and accurate determination of the pressure and vibration characteristics of any similar system.

Summary and Conclusions

For pumping crude oil at the lower pressures (up to 900 psi), sufficient information is available for proper design of a piping system to avoid undue mechanical vibration. The design of low-speed duplex and triplex horizontal pumps is satisfactory at the lower pressures. The use of equalizer lines for two or more units on one system is a satisfactory practical solution at lower pump speeds and pressures.

The design of high-speed short-stroke pumps (175 rpm and greater) to handle large quantities of crude is not sufficiently worked out to avoid probability of breakage difficulties in the pump or piping. These factors become more important at the higher pressures now used.

The varying-grade crudes pumped, the use of a smaller number of large-capacity pumps, and the higher design pressures, encountered on a modern pipe-line system complicate the proper design of a high-speed multicylinder pump. Solutions to these difficulties are being worked out in actual field installations at the present time. It has not been found practical for the pump companies to set up laboratory installations for the size units used in a modern pipe-line pumping station.

The higher frequencies produced by the pump as pressure surges do not complicate the problem of piping design unduly. Avoidance of sharp discontinuities in the piping system and concentrations of mass or unsupported piping sections should result in a satisfactory piping system. The use of air chambers is not justified at surge frequencies above 20 cps.

Mechanical vibration transmitted from the engine or pump to the piping system is considered as a separate problem from the one covered in this paper.

Our general requirements for suitable equipment for pressure-surge and vibration measurements have been fairly well established. There should be no undue difficulties in setting up such a system with the exception of locating a pressure gage with suitable characteristics.

Several systems on the market for determining pressure variations have been studied for our use in this work. These have consisted mostly of systems designed primarily for engine-characteristics measurements and are not believed suitable for our application.

Acknowledgments

The author wishes to express his debt of gratitude to all those who have been so helpful in assisting in the preparation of this paper. Representatives of the Aldrich Pump Company, Worthington Pump and Machinery Corporation, National Transit, and Cooper-Bessemer Corporation have furnished valuable data and sources of information.

Much credit is also due to the members of the engineering staff and the departments of the Stanolind Pipe Line Company for their constructive criticisms and ideas offered, particularly Mr. T. R. Aude, Mr. J. R. Polston, Mr. A. H. Newberg, Mr. H. L. Drake, Mr. R. P. Lennart, and Mr. C. F. Gearhart. Mr. J. B. Eisler and Mr. T. Gilmartin of the Stanolind Oil and Gas Company have also been generous with their time in helping to formulate some of the conclusions herein expressed.

Appendix

Speed of Pressure Wave in a Pipe

Symbols Used

S_u = longitudinal stress, psi
S_v = circumferential stress, psi
P = pressure, psi
r = pipe radius, in.
t = pipe thickness, in.
μ = Poisson's ratio = 0.30 for steel
α = $1/\mu$ = 3.33 for steel
x = length of pipe, in.
y = circumference of pipe, in.

e_u = longitudinal strain
e_v = circumferential strain
E = modulus of elasticity, psi (steel = 30×10^6)
K = bulk modulus of liquid, psi (37 deg API oil at 60 F = 200,000)
V = original volume
ΔV_P = change in volume of pipe (expansion)
ΔV_L = change in volume of liquid (compression)
ΔV = total relative change in volume due to expansion of pipe and compression of liquid
Δr = change in radius
Δx = change in length
C = speed of sound or pressure wave in combination of liquid and pipe
J = effective modulus of elasticity of pipe and liquid, accounting for compressibility of liquid, as well as circumferential and longitudinal expansion of pipe, psi
M = mass of oil, pcf
g = acceleration of gravity = 32.2 fps per sec

In the case of a pipe carrying oil, the velocity of a pressure (or sound) wave is modified by the elasticity of the oil and the steel. For the pipe

$$S_u = \frac{Pr}{2t};\ S_v = \frac{Pr}{t} \quad \text{[1]}^{8,9}$$

From Poisson

$$e_u = \frac{\Delta x}{x} = \frac{S_u}{E} - \frac{S_v}{\alpha E} = \frac{Pr}{2tE}\left(1 - \frac{2}{\alpha}\right) \quad [2]$$

$$e_v = \frac{\Delta y}{y} = \frac{\Delta r}{r} = \frac{S_v}{E} - \frac{S_u}{E} = \frac{Pr}{2tE}\left(2 - \frac{1}{\alpha}\right) \quad [3]^{8,9}$$

Since

$$\frac{\Delta y}{y} = \frac{\Delta(2\pi r)}{2\pi r} = \frac{\Delta r}{r}$$

$$V + \Delta V_p = (x + \Delta x)\,(r + \Delta r)^2\pi$$

By neglecting second and third-order derivatives. This may be expressed

$$\Delta V_p = 2\pi r x\,\Delta r + \pi r^2\,\Delta x$$

$$\Delta V_p = \pi r^2 x \cdot \left(\frac{2\,\Delta r}{r} + \frac{\Delta x}{x}\right) \quad [4]$$

but

$$\pi r^2 x = V$$

Then substituting Equations [2] and [3]

$$\Delta V_p = \frac{Pr}{2tE}\,V \cdot \left(5 - \frac{4}{\alpha}\right) \text{ for the pipe} \quad [5]$$

By definition

$$\frac{\Delta V_L}{V} = \frac{P}{K} \quad \text{or} \quad \Delta V_L = \frac{PV}{K} \text{ for the liquid} \quad [6]$$

$$\Delta V = \Delta V_P + \Delta V_L = \frac{PV}{J}$$

$$\frac{PV}{J} = \frac{PrV}{2tE}\left(5 - \frac{4}{\alpha}\right) + \frac{PV}{K}$$

$$\frac{1}{J} = \frac{r}{2tE}\left(5 - \frac{4}{\alpha}\right) + \frac{1}{K} \quad [7]$$

$$C = \frac{12}{\sqrt{\frac{M}{g}\left(\frac{1}{J}\right)}} = \frac{12}{\sqrt{\frac{M}{gK}\left[1 + \frac{rK}{2tE} \cdot \left(5 - \frac{4}{\alpha}\right)\right]}} \quad [8]$$

for crude oil in steel pipe, assuming $K = 200{,}000$ psi; $E = 30 \times 10^6$ psi; $M = 53$ pcf; $\alpha = 3.33$, Equation [8] reduces to

$$C = \frac{4180}{\sqrt{1 + \frac{0.0127r}{t}}} \text{ fps} \quad [9]$$

Joukovsky has developed a formula which does not account for the longitudinal pipe movement. Since this small factor may be neglected, we may also use

$$C = \frac{12}{\sqrt{\frac{M}{gK}\left(1 + \frac{2rK}{tE}\right)}} \quad (10)^{10,11}$$

Or for the same valves as in Equation [9]

$$C = \frac{4180}{\sqrt{1 + 0.0133\,\frac{r}{t}}} \text{ fps} \quad [11]$$

Notes on Air-Chamber Calculations

Since the advent of the use of equalizer lines in our system, it has not been found necessary to use air chambers or alleviators on the discharge side of reciprocating pumps. In order to provide for constant flow of the oil mass by a cushioning or smoothing effect on the suction side of the slower-speed duplex and triplex pumps, particularly where the installation is not a flooded or positive-pressure suction, it is felt highly desirable to use air chambers at the pump.

However, where the frequency of volumetric change in the pump cycle becomes comparatively high (an estimated 25 cps or more), and the volumetric change in the resultant flow curve of the pump is small, it is not felt necessary to introduce an air chamber. Here it is believed that (1) the inertia of the oil mass is such that it will not follow these changes in volume; and (2) the amount of volume variation in a properly operating pump of 5 or more cylinders at a speed of 175 rpm or greater is insufficient to justify the use of an air chamber.

Calculations for sizing air chambers following the equations developed by Prof. E. H. Kennard,[12] and those which we have developed on the basis of sizing the air chamber so that the air and oil act as a resonant spring-mass system, have borne out these observations. This latter derivation is not included here because of its narrow usage in practice. A few field observations have also been made to substantiate this view further.

[8] "Advanced Mechanics of Materials," by Glenn Murphy, McGraw-Hill Book Company, Inc., New York, N. Y., 1946.

[9] "Hydraulics of Pipe Lines," by W. F. Durand, D. Van Nostrand Company, Inc., New York, N. Y., 1921.

[10] "A Treatise on Applied Hydraulics," by Herbert Addison, third edition, John Wiley & Sons, Inc., New York, N. Y., 1945.

[11] "Water Hammer," by Miss O. Simin (digest of Prof. N. Joukovsky's work), Proceedings of the American Water Works Association, vol. 24, 1904, pp. 341–424.

"Water-Hammer Problems in Connection With the Design of Hydroelectric Plants," by E. F. Strowger, Trans. ASME, vol. 67, 1945, pp. 377–388.

[12] "The Minimum Size of Air Chamber Required to Prevent the Inception of Elastic Surges in Pipe Lines From Reciprocating Pumps," by E. H. Kennard, *Sibley Journal of Engineering*, vol. 43, May, 1929, pp. 187–188, 222 and 224.

Continental and American Gas-Turbine and Compressor Calculation Methods Compared

By P. F. MARTINUZZI,[1] ITHACA, N. Y.

During a recent tour of the United States, it became evident to the author that there is no standard American method for gas-turbine or for axial-compressor calculations. Apparently, each firm goes its own way without much regard for precedents, competitors, or textbooks. In view of the wide diffusion that British data had in this country during the war, it might have been supposed that the Howell (1)[2] method would have been generally used. However, although the actual data given by Howell are considered very reliable, his calculation method is not much used. In practice, calculations seem mostly to be influenced by Keller (2) and, to a much lesser degree, by Traupel (3). The same situation can be noted as regards gas-turbine-cycle calculations; the paper by Soderberg and Smith (4) is generally well known, but no standard system seems to prevail. To simplify matters, it will be assumed that general American practice can be represented by the methods given in Professor Zucrow's (5) recent book and in the papers by Ponomareff (6), and Howell (1), as regards axial compressors.

Nomenclature

THE following nomenclature is used in this part of the paper. It represents average continental practice; but symbols are not completely standardized even there. Dimensions are metric, but wide use of dimensionless coefficients makes application to American standards easy.

a = velocity of sound, mps
b = radial length of blade, m
c_A = lift coefficient
c_w = drag coefficient
c = absolute velocity (located according to subscript), mps
c_m = axial component of absolute velocity, mps
c_u = tangential component of absolute velocity, mps
d = diameter, m
l = blade chord length, m
M = Mach number
n = speed, rpm
p = pressure, kg per sq m
r = radius, m
R = degree of reaction
t = blade pitch, m
u = peripheral velocity, mps
w = relative velocity (in rotor), mps
w_u = tangential component of relative velocity, mps
w_∞ = mean relative velocity in rotor (wind velocity), mps
β = angle between relative velocity and tangential direction
γ = angle between absolute velocity and tangential direction
δ = glide angle in stator
ϵ = glide angle in rotor
η = efficiency (stage or part of stage)
τ = ratio c_u/u (vectorial)
τ_s = $\tau_1 + \tau_2$ (vectorial)
τ_d = $\tau_1 - \tau_2$ (vectorial)
φ = dimensionless flow coefficient = c_m/u
ψ = dimensionless pressure coefficient = $\dfrac{\Delta p}{\rho u^2/2}$

Subscripts: 0, at pitch diameter; s, at tip diameter; N, at hub diameter
Subscripts: 1, at rotor inlet (or stator outlet); 2, at rotor outlet (or stator inlet)
Superscripts: $'$, in the stator; $''$, in the rotor

Basis of Continental Theory

Continental industrial gas turbines run at lower temperatures than those usual in America. Maximum temperatures of 1120 F and 1200 F are employed in the power plants under construction; and 1300 F is the limit for future projects. This is due to difficulties in obtaining good heat-resisting alloys and to the desire for reliability; also, closed and semiclosed cycles are less suited to high temperatures. As very high thermal efficiencies are wanted, owing to the high prices of fuels, particular care has been given on the Continent to high adiabatic efficiencies for compressors and turbines. Ease of manufacture and compactness are secondary considerations; high efficiency is paramount. Consequently, axial compressors are calculated according to a theory which tends to give very high efficiencies; it is also used for turbines. The basic principles on which it is founded and an outline of the calculation method will be given.

Considering one stage of the compressor, the following simplifying assumptions are made:

(*a*) Density variations so small that density can be considered constant.
(*b*) Constant axial velocity c_m at all diameters.
(*c*) Constant cross section throughout the stage.
(*d*) Potential motion.
(*e*) Lines of flow moving axially along cylindrical surfaces.

These assumptions define what is generally called a "normal stage"; more general assumptions are considered by Traupel (3), but the theory actually used is based on the normal stage. Starting from these assumptions, and considering a blade element between two cylindrical bounding surfaces, a blade-element theory is developed in the known manner, as given, for instance, by Professor Zucrow.[3] The aerodynamic forces acting on the blade element are determined; the ideal pressure rise is calculated. Considering real flow with losses, use is made of the cascade polar and of the lift and drag coefficients c_A and c_w; these, or the glide angles c_w/c_A (called δ in the stator and ϵ in the rotor), are introduced, and the effective pressure rise is calculated; the cascade characteristics, blade chord length l, and blade pitch dis-

[1] Professor of Mechanical Engineering, Cornell University; formerly Head of Gas Turbine Section, Italian National Research Council, Rome, Italy. Mem. ASME.

[2] Numbers in parentheses refer to Bibliography at end of the paper.

Calvin W. Rice Lecture for 1948, contributed by the Gas Turbine Power Division and presented at the Semi-Annual Meeting, Milwaukee, Wis., May 30–June 5, 1948, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—SA-62.

[3] Bibliography, reference (5), chapt. 9.

tance t are brought in by means of the concept of circulation. Full developments and proofs are given elsewhere by the author.[4]

Dimensionless Quantities

The continental method is rendered very flexible by the extended use of dimensionless quantities and dimensionless velocity diagrams. The tangential components of the absolute velocities, c_{u1} and c_{u2}, are given in the form of ratio to the peripheral velocity u, and so is the axial component c_m

$$\tau_1 = \frac{c_{u1}}{u} \quad [1]$$

$$\tau_2 = \frac{c_{u2}}{u} \quad [2]$$

$$\varphi = \frac{c_m}{u} \quad [3]$$

$$\tau_s = \tau_1 + \tau_2 \quad [4]$$

$$\tau_d = \tau_1 - \tau_2 \quad [5]$$

The quantities τ are vectors; they are considered positive when they have the opposite direction of the peripheral velocity (this conventional assumption is made to show that the machine is a compressor and not a turbine). Using these quantities, it is possible to draw dimensionless diagrams, Figs. 1 and 2, in which u is considered unity; they are very useful.

The quantity φ is called the flow coefficient because, for a given section and peripheral velocity, it indicates the volume flow through the compressor. The dimensionless pressure coefficient, which compares the pressure increase in the stage to the pressure head corresponding to the peripheral velocity, is

$$\psi = \frac{\Delta_p}{\rho u^2/2} \quad [6]$$

φ, ψ and another dimensionless coefficient σ, now fallen into disuse on the Continent, were introduced by Keller (2). With these dimensionless quantities, calculations become very simple. It can be shown that the theoretical pressure coefficient is

$$\psi_t = 2\tau_d \quad [7]$$

Given the glide angles δ and ϵ in stator and rotor, the efficiency of the blade element becomes

$$\eta = \frac{4\varphi - (\epsilon + \delta)(4\varphi + 2\tau_s + \tau_s^2)}{4\varphi + 4\epsilon + 2(\epsilon - \delta)\tau_s} = \frac{\psi}{2\tau_d} \quad [8]$$

Though apparently complicated, this formula is easy to use; δ and ϵ are usually known (they vary generally between 0.02 and 0.04); and τ_s and τ_d assume simple values, once the blading type is given, as will be seen.

The method given by Howell (1) differs considerably in appearance from the continental method, as Howell uses the deflection angles as main parameters; but it can be shown that the difference is formal; it is obvious, for instance, that for a given φ, τ_d is a measure of the deflection.

Degree of Reaction

The degree of reaction is the ratio Δ_p''/Δ_p between the pressure increase Δ_p'' obtained in the rotor and the total pressure rise in the stage. It must be remembered that, up to now, we have considered only one blade section at a certain diameter; the degree of reaction R pertains to that section, and varies, in general, from section to section along the diameter. If the slight difference caused by the drag losses is neglected, that is, if the ideal pressure rises in rotor and stage are considered, instead of the real ones, it can be shown that

$$R = 1 + \tau_s/2 \quad [9]$$

It can easily be shown that, in the dimensionless velocity diagrams, Fig. 1(a), and Fig. 2(a), R is measured by the tangential component of the mean wind velocity w_∞. As will be seen, R is an important quantity and it is useful to have an easy way of determining it.

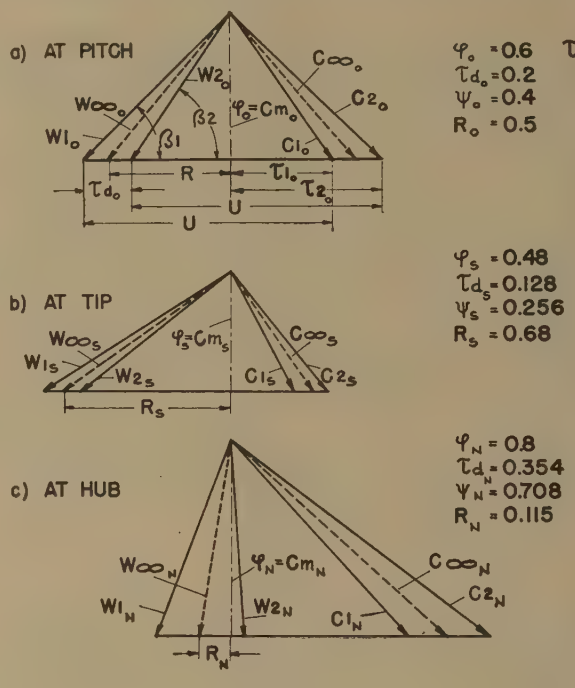


Fig. 1

Types of Blading

Up to now, no assumptions were made regarding either the type of blading or the position along the diameter of the blade element considered. On the Continent, types of blading are defined by the form of the dimensionless velocity triangles at the "pitch diameter." Pitch-diameter d_0 is generally, but not necessarily, halfway between hub diameter d_N and tip diameter d_s. The types of blading used more generally on the Continent for gas-turbine compressors are as follows:

(a) The spinless rotor-outlet type, Fig. 2(a).

(b) The symmetrical, or congruent, or 0.5 reaction type, Fig. 1(a).

Less frequently used in gas turbines are the following:

(c) The spinless rotor-inlet type, Fig. 3(a)

(d) The symmetrical stator type, Fig. 3(b).

It must be noted that the very widely used (b) type is not equivalent to what is called in America, for instance, by Ponomareff (6), symmetric or constant-reaction type. The (b) type continental blading has 0.5 reaction and symmetrical diagram only at the pitch diameter, and has free vortex twist along the radius. In this paper the American type symmetrical blading will always be called "constant-reaction blading." If Ponomareff's paper (6) is not read carefully, it might appear that only types (a) and (c), called in that paper "vortex-inlet flow" and

[4] Bibliography, reference (7).

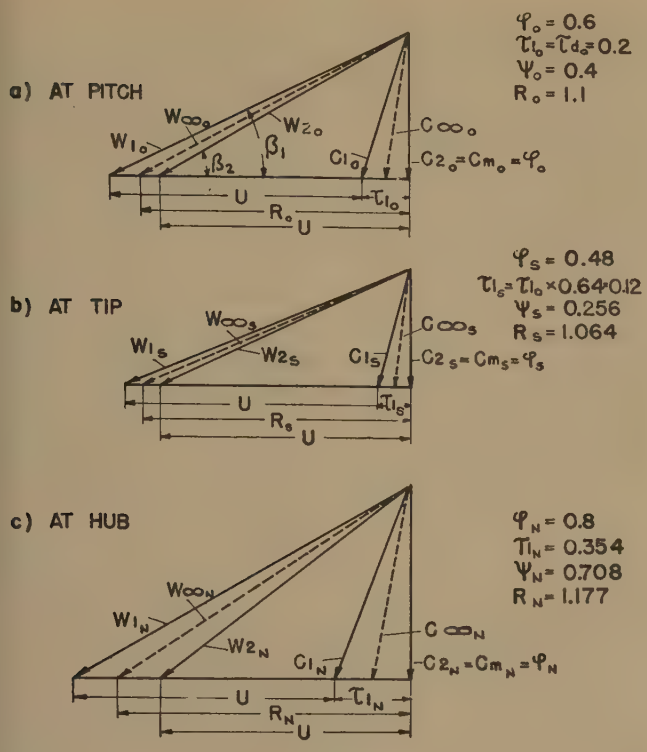


FIG. 2

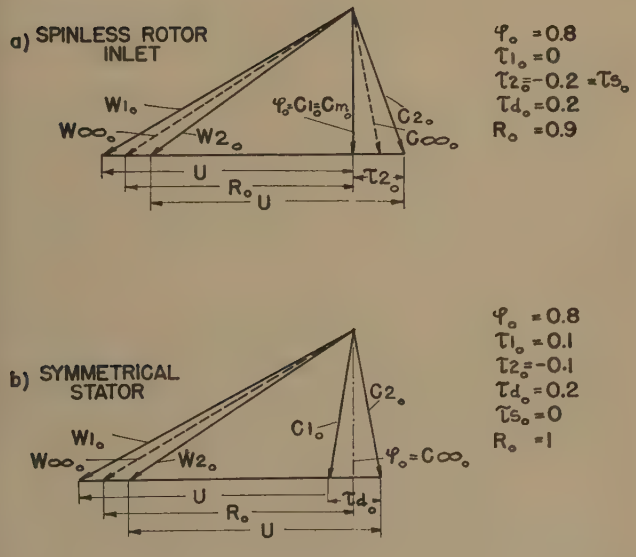


FIG. 3

"axial inlet flow" are free-vortex blades; actually, on the Continent all four blade types are of the free-vortex variety.

The advantages of the dimensionless quantities become apparent when these usual types of blading are considered. Thus it can be shown that, for blade type (*a*) (spinless rotor outlet)

$$\tau_{s0} = \tau_{d0} = \tau_{10} > 0,\ \tau_{20} = 0 \quad [10]$$

consequently

$$\psi_0 = 2\eta\tau_{10} \quad [11]$$

$$R = 1 + \tau_{10}/2 > 1 \text{ (Fig. 2}a\text{)} \quad [12]$$

For type (*b*) (symmetrical or 0.5 reaction), $\tau_{d\,0}$ can be chosen at will

$$\tau_{s0} = -1 \quad [13]$$

(because w_2 is symmetrical to c_1; the sum of the tangential projections of w_2 and c_2 is obviously u, or unity; and the tangential projections of c_1 and c_2 are directed like u; and therefore are negative according to the convention made). Hence

$$\psi_0 = 2\eta\tau_{d0} \quad [14]$$

$$R = 1 - {}^1/_2 = 0.5 \text{ (Fig. 1}a\text{)} \quad [15]$$

For type (*c*) (spinless rotor inlet)

$$\tau_{10} = 0;\ \tau_{20} < 0;\ \tau_{s0} = \tau_{20} = -|\tau_{20}|;\ \tau_{d0} = -\tau_{20} = +|\tau_{20}| \quad [16]$$

$$\psi_0 = 2\eta|\tau_{20}| \quad [17]$$

$$R = 1 - |\tau_{20}/2| < 1 \text{ (Fig. 3}a\text{)} \quad [18]$$

For type (*d*) (symmetrical stator)

$$\tau_{10} = -\tau_{20};\ \tau_{s0} = 0;\ \tau_{d0} = 2\tau_{10} \quad [19]$$

$$\psi_0 = 4\eta\tau_{10} \quad [20]$$

$$R = 1 \quad [21]$$

Conditions Along Radius

The continental theory is based upon the assumption that the total pressure rise in the stage must be the same at all diameters along the radius, so as to avoid radial pressure differences which might cause undesirable inflections of the flow lines and longitudinal whirls, with consequent loss of efficiency. Particular importance is given to radial pressure equilibrium in the axial gaps before and after the rotor. It can be shown that these conditions can be satisfied if there is constant circulation along the blade and if the fluid motion in the gaps is that of a free vortex, provided the axial velocity c_m is constant at all diameters, as assumed before. These conditions are followed rigidly in continental calculations. We shall see later whether this is always advisable. Assuming constant circulation, free-vortex distribution and constant axial velocity, it can be shown that the following expressions are valid; quantities without subscripts being blade-element conditions at a diameter d chosen at will, and those with subscript 0 being at pitch diameter d_0

$$\varphi = \varphi_0 \frac{d_0}{d} \quad [22]$$

$$\tau_s = \tau_{s0}\left(\frac{d_0}{d}\right)^2;\ \tau_d = \tau_{d0}\left(\frac{d_0}{d}\right)^2 \quad [23]$$

$$\psi = \psi_0\left(\frac{d_0}{d}\right)^2 \quad [24]$$

From Equation [9] it follows that

$$R - 1 = \tau_s/2 \quad [25]$$

hence

$$R - 1 = (R_0 - 1)\left(\frac{d_0}{d}\right)^2 \quad [26]$$

$(R - 1)$ is the pressure rise obtained in the stator; it varies with the reciprocal of the diameter ratio, and the variation increases with the absolute value of $(R_0 - 1)$.

The cascade polar can be introduced considering the condition of constant circulation at all diameters. It follows

$$c_A \frac{l}{t} = \frac{2\tau_d}{\sqrt{R^2 + \varphi^2}} \quad [27]$$

at all diameters.

As the blade pitch t decreases proportionately to the diameter, c_A must increase at the hub diameter because Equation [23] shows that τ_d increases strongly there; this increase of c_A can be obtained either by increasing the incidence of the blade profile relative to w_∞, or by using a more cambered profile at the hub. But the increase in c_A is limited by the danger of stalling (a clear definition of stalling in cascades is given by Howell) (1). Making the blade chord l larger at the hub helps a little, but not much.

Obviously, conditions at the hub are critical points in the calculation, owing to cascade limitations. This can be expressed by a formula setting a maximum limit Z to the cascade quantities

$$c_A \frac{l}{t} \leqq Z \quad [28]$$

where Z is determined by cascade tests and practical experience. Keller (2), for the use of single-wing data instead of cascades, gives $Z = 0.7$. The actual value of Z used nowadays varies from firm to firm, depending on cascade and multistage-compressor tests, and is generally kept secret. It is probable that, with appropriate profiles and angles, good efficiencies can be obtained with $Z = 0.9$. If condition, Equation [28], is satisfied at the hub, obviously, it is satisfied everywhere else. Combining Equations [27] and [28], and remembering that the subscript N shows hub conditions

$$\frac{\tau_{dN}}{\sqrt{R_N^2 + \varphi_N^2}} \leqq Z \quad [29]$$

It can be shown that, using Equations [22] and [23] and the diameter ratio d_0/d_N, in Equation [29] can be written

$$\frac{2\tau_{d\,0}\left(\frac{d_0}{d_N}\right)}{\sqrt{\varphi_0^2 + \tau_{s0} + \left(\frac{d_N}{d_0}\right)^2 + \left(\frac{d_0}{d_N}\right)^2 \frac{\tau_{s0}^2}{4}}} \leqq Z \quad [30]$$

This expression is fundamental; given the diameter ratio, it allows the maximum possible τ_{d0} (and hence the maximum possible pressure rise per stage ψ_0) to be calculated, or vice versa. It looks complicated, but actually lends itself very easily to calculation; for the usual types of blading, expressions such as Equations [10] and [13] simplify matters considerably, as will be seen later.

At the blade tip, there are no difficulties with stalling, but limiting conditions are set by compressibility. When the relative wind velocity w_∞ approaches the speed of sound, efficiency falls rapidly. The actual velocities round the blade profile are greater than w_∞, owing to circulation; the ratio between the two can be improved if appropriate flat profiles are chosen. It is necessary to limit w_∞ at the tip to a fraction of the prevailing sound velocity a; this fraction or ratio $\mathbf{M}$ (the maximum allowable Mach number) depends upon the profile and its cascade characteristics. At all diameters, as can be seen from the definition given for the degree of reaction, the value of w is given by

$$\frac{w_\infty}{u} = \sqrt{R^2 + \varphi^2} \quad [31]$$

Introducing the usual dimensionless quantities and the diameter ratio d_0/d_s between pitch and tip diameters, this can be written

$$u_0 \sqrt{\varphi_0^2 + \tau_{s0} + \left(\frac{d_s}{d_0}\right)^2 + \left(\frac{d_0}{d_s}\right)^2 \frac{\tau_{s0}^2}{4}} \leqq \mathbf{M}a \quad [32]$$

This is the second fundamental equation; once the cascade type and its allowable Mach number are chosen, the maximum permissible peripheral speed for a given diameter ratio can be determined or vice versa. $\mathbf{M}$ is generally not higher than 0.7 to 0.75 for high-efficiency compressors. Recent continental tests, however, seem to indicate that if appropriate profiles are chosen good efficiencies can be obtained even with $\mathbf{M} = 0.8$ to 0.9.

Compressor Calculations

Equations [30] and [32] are the two basic expressions used in practice when calculating axial compressors by the continental method. Given the type of cascade and the blading efficiency η (always very high, $\eta = 0.92$ — 0.95; with a little practice, it can be estimated) or stage efficiency η_s (also this can be estimated), for a given diameter ratio, Equation [30] gives τ_{d0} and ψ_0, and Equation [32] gives u; consequently, the stage pressure rise

$$\Delta p = \rho \psi_0 \frac{u^2}{2} \quad [33]$$

can be calculated. For usual types of blading, these formulas become simpler; thus for the widely used 0.5 reaction blading, $\tau_{s0} = -1$, which makes computing much easier.

The limitations set by Equations [30] and [32] are very important and must be respected. However, it is sometimes necessary to sail close to the wind, owing to weight or space considerations. In that case, it is better to choose too large a Mach number $\mathbf{M}$ in Equation [32] rather than too high a value of Z in Equation [30], for the following reasons:

1 An increase of $\mathbf{M}$ affects u_0, which appears in the square in Equation [33], while an increase of Z affects ψ_0, which appears in the first power.

2 Both expressions are critical for the first (lowest pressure) stage, because there specific volumes are largest, blades longest and diameter ratio widest; but Equation [32] is doubly critical at low pressure, because it contains the velocity of sound a, which increases with the pressure (or rather, with the square root of the absolute temperature); consequently, passing from lower to higher stages, conditions will improve more rapidly at the tip than at the hub.

3 When the flow decreases, that is, when φ_0 becomes smaller, conditions rapidly get worse at the hub (as shown by Equation [29]), while they improve at the tip, as shown by Equation [31]; and vice versa when the flow is increased; but good conditions at increased flow are less important than at reduced flow, when pumping and other troubles are liable to occur.

Gas-Turbine-Compressor Design

The aim of the gas-turbine designer is to build an axial compressor with as few stages and as small a bulk and cost as possible. This is imperative in the case of aircraft turbines, to reduce weight and cross section. But even in industrial and marine turbines, a low number of stages is desirable, because the flow, in passing from stage to stage, becomes progressively more uneven and disturbed; if too many stages are used, the efficiency of the later stages falls because the air is already full of eddies when entering them. On the Continent, it is considered advisable not to have more than 12–15 stages per compressor; if more are needed, then it is better to split up the compressor and use an intercooler. Apart from the gas-turbine-cycle advantages (which are notable only if a heat exchanger is used), an intercooler, by slowing down the air and then accelerating it again, makes the flow smooth again and allows the high-pressure

stages to work at good efficiency. Also, it is then often possible to change the pitch diameter, thus allowing the use of more suitable blade lengths; or even, when two turbines are used, to vary the rotational speed of the two compressors.

This question of maximum number of stages becomes more important when permissible turbine-inlet temperatures increase and require higher compression ratios for maximum cycle efficiency. Ratios of 6 to 1 are already in use; and propeller-turbine units will have to go much higher, 10–12 to 1, if reasonable fuel consumption is to be obtained without a heat exchanger. Some possible solutions of the problem will be indicated later in this paper.

The requisites for reduced over-all diameter and reduced number of stages are contradictory. Once the type of blading is chosen, the smallest number of stages will be obtained when u_0 will be highest, u_0 being the peripheral velocity at pitch (that is, mean) diameter. On the other hand, the Mach-number limitations affect the tip speed; consequently, it pays to reduce the blade length, i.e., the diameter ratio, so as to bring the pitch speed as near the tip speed as possible, as proved by Expression [32]. Besides, a diameter ratio near unity increases $\tau_{d\,0}$, as shown by Equation [30], thus further increasing the pressure obtainable per stage. On the other hand, given u_0 and the axial velocity c_m (which cannot be increased beyond a certain limit to contain friction and diffuser losses within reason; on the Continent, c varies between 300 and 500 fps, in England and in the United States higher values, up to 700 fps, are used), the flow cross section is determined. If the blade length b is decreased, then the required cross section must be obtained by increasing the pitch diameter d and decreasing correspondingly the rotational speed. In this way, the over-all diameter of the compressor is increased. Consequently, the smallest number of stages and smallest over-all diameter are contradictory requirements.

A reduction in blade height certainly causes a notable reduction in number of stages; for instance, everything else being the same, a reduction in the blade length: pitch-diameter ratio b/d_0 (which is the normal way of gaging blade length on the Continent) from $^1/_5$ to $^1/_9$ would reduce the number of stages from 10 to 6, for the same compression ratio. Besides, stresses and vibrations are smaller in shorter blades. Hence in America and in England, reasonably short blades seem to be favored in industrial and even more in aviation turbines. On the other hand, on the continent longer blades are used, even though they require a larger number of stages.

There are several reasons for this. Short blades give lower stage and compressor efficiencies; and on the Continent, high efficiency is always the first consideration. A short blade has a low aspect ratio; and, although it is generally assumed that the rotor and stator cylindrical boundary surfaces constitute effective end vortex shields, this is not actually the case, especially on the side of the clearance. There are end vortex losses, and these become smaller if the aspect ratio is increased. However, apart from the aspect ratio, the absolute length of the blade counts too; the thickness of the boundary layers on the rotor outer and stator inner cylindrical surfaces is roughly independent of blade length. Consequently, the longer the blade, the smaller the noxious influence of the boundary layers.

Usually, on the Continent, blades are never made shorter than 40 mm (about 1.6 in.) on the shortest (high-pressure) stage. If necessary, the pitch-diameter line is decreased from stage to stage, so that the compressor is conical (with the smaller diameter at the high-pressure end), to prevent the blades from being too short. Calculations are set up starting from a diameter ratio, to be used in Equations [30] and [32], corresponding to blade length/pitch diameter ratio b/d_0 of $^1/_5$ and even, if the volumes are very large, $^1/_4$.

In this way the minimum compressor outer diameter is obtained, given c_m; also, the rotor diameter becomes as small as possible, diminishing cost and weight. It must be noted that the longer blades and greater number of stages with relatively lower peripheral speeds imply a lower blade loading. This sometimes allows a decrease in blade chord, increasing the aspect ratio, improving the cascade characteristics, and actually counteracting the increase in length due to the greater number of stages. Also continental industrial gas turbines often must deal with larger volumes of air than their American counterparts, and the lower temperatures further increase the weight of air per horsepower. Consequently, long blades are necessary on many continental compressors simply to cope with the flow.

The actual procedure in calculating axial compressors is as follows: Axial velocities, stage and compressor efficiency, M and Z are assumed, based upon known examples; the flow cross sections before the first and after the last stage are calculated; the compression ratio desired being known, sound velocity before the first stage also is known. Choosing tentatively a b/d_0 ratio of $^1/_5$, the maximum values of $\tau_{d\,0}$ and u_0 are then given by Equations [30] and [32] for the first stage. Consequently, the pressure rise Δp, d_0, and the rpm can be calculated. These values of d_0 and rpm are then applied to the last stage, and the blade length b is then calculated from the flow cross section. If b is large enough (i.e., greater than 1.6 in.), then the pressure rise for the last stage will be calculated (it is greater than for the first: density is greater and the diameter ratio nearer unity permits a larger $\tau_{d\,0}$). The mean value of Δp for the first and last stages will then give an approximate idea of the number of stages. If, however, as often is the case, b for the last stage is insufficient, then either c_m is decreased at the high-pressure end (this is always advisable, as it allows the whole compressor length to be used as a partial diffuser, thus reducing leaving losses), or d_0 is reduced at the high-pressure end, or both.

After repeated attempts, in which some of the assumptions are modified if necessary, a reasonable compromise is reached. The resulting data, and particularly the rpm, are then matched with those of the turbine which drives the compressor. After further fumblings, another compromise is reached. Then the whole combination, compressor and turbine, is calculated anew, introducing more exact data, calculating the efficiencies instead of assuming them, etc.; the exact number of stages, for instance, must be determined by getting the pressure rise stage by stage; if not, mistakes of one or two stages are easily made.

Calculating gas turbines is a long job, because an alteration in any element of the cycle reflects on everything else, and often makes a complete recalculation necessary. On the Continent, the number of engineers and designers used on each gas-turbine project is enormous; but practice has shown that much money and many needless experiments are saved if a well-trained and co-operative calculation and design team is available.

Blading Types Compared

Up to now, nothing has been said about blading types. If the lowest possible number of stages is desired, however, a glance at Fig. 1(a) and Fig. 2(a) will show that the 0.5 reaction type is superior to the spinless rotor-outlet type. It has been seen that the limiting factor for stage pressure rise is the Mach number of the mean wind velocity w_∞. The figures show that in the spinless outlet type, w_∞ must necessarily be larger than the peripheral velocity u_0; while in the 0.5 reaction type, w_∞ must be larger than one half of u_0; this is shown clearly by expression, Equation [31], because, for the spinless outlet type, $R_0 > 1$, while for the 0.5 reaction type of course $R_0 = 0.5$. Consequently, given the Mach number and the upper limit of w_∞, u_0 is much higher for the 0.5 type than for the spinless type; and the

pressure rise is proportional to u_0^2. In actual fact, however, the advantage of the 0.5 reaction type is not as big as it appears at first sight, for the following reasons:

(*a*) The limiting Mach number affects conditions at the blade tip, not at the pitch diameter. In a free-vortex 0.5 reaction blading (and continental theory considers only free-vortex blades), the degree of reaction at the tip is much greater than 0.5; as shown by Fig. 1(*b*) (the figures give an extreme case, the b/d ratio being $^1/_4$), it is $R_s = 0.68$, representing a 36 per cent increase over pitch-diameter conditions.* On the other hand, in the case of the free-vortex spinless outlet blading, Fig. 2(*b*) shows that there is actually a decrease in the degree of reaction from pitch to tip, $R_0 = 1.1$ and $R_s = 1.064$. (The reason for this variation of R can be gathered from Expression [26], $R - 1$ being positive for the spinless outlet type and negative for the 0.5 reaction type.). Hence the difference as regards u_0 is smaller than would appear from pitch-diameter conditions.

(*b*) Beside conditions at the tip, limiting conditions at the hub must be considered too, and there, the spinless outlet type scores heavily; the fact that, for a given u_0, it has a higher w_∞ is an advantage at the hub, as this reduces the incidence angles required for a given load and allows much higher values of τ_d to be used without danger of stalling. This is shown clearly by Equation [29]; an increase in R allows a much larger τ_{d0} for a given Z. Again, it must be noted that the blade twist favors the spinless outlet type; in it, the decree of reaction is slightly larger ($R_N = 1.177$) at the hub than at pitch diameter; while in the 0.5 reaction type, $R_N = 0.115$ at the hub is very much smaller than at the pitch diameter, as shown in Fig. 2(*c*) and Fig. 1(*c*), which give the triangles at the hub. The reason again is given by Equation [26]. Consequently, the usual permissible values are $\tau_{d0} = 0.18 - 0.26$ for the 0.5 reaction type; $\tau_{d0} = 0.34 - 0.44$ for the spinless outlet type.

Even with these limitations, the advantage of the 0.5 reaction type is considerable. It can also be seen that this type lends itself to larger values of φ_0 and higher axial velocities than the spinless outlet type, because an increase of φ_0 in Equations [29] and [31] is more than compensated by the decrease of R_0, and because, for a given φ_0, the larger u_0 implies a larger c_m. Consequently, the 0.5 reaction-type compressor is not only shorter but also smaller in cross section. It is particularly useful in aviation work.

It must be noted, however, that the spinless outlet type is slightly more efficient. The fact alone of having a spinless outlet reduces the leaving losses, though these are relatively more noxious in ventilators or compressors with few stages than in gas-turbine compressors. In practice, however, the efficiency of a multistage compressor of the spinless outlet type is higher than that of other types for the following very important reasons:

In the spinless outlet type, the degree of reaction is greater than 1; this means of course that the pressure actually drops in the stator, and then is built up again in the rotor, which consequently must deliver a pressure greater than the stage pressure rise required. At first sight, this seems a rather foolish thing to do, but in practice it is a great advantage. The pressure drops in the stator of course because in it the flow is accelerated, Fig. 2(*a*), where stator-entry velocity $c_2 < c_1$, stator-exit velocity). An acceleration of the flow has a remarkably smoothing effect, particularly on the boundary layer. This acceleration not only allows the flow to pass through the stator without trouble, but even cancels or reduces eddies and disturbances caused by the preceding stages. That is why the spinless outlet type allows more stages to be used without trouble than any other type. Of the four types of continental blading described, the spinless outlet is the only one which enjoys this smoothing property. Consequently it is much superior in this respect not only to the 0.5 reaction type, where the flow is decelerated equally in stator and rotor, but also to the spinless rotor-inlet type, Fig. 3(*a*), where there is a small deceleration in the stator, and to the symmetrical stator type, Fig. 3(*b*), where the velocity in the stator should be constant, but is in practice slowed down by friction. Only a definite acceleration has a definite curing effect for boundary-layer trouble.

Coming back to our comparison between spinless outlet and 0.5 reaction types, there is no doubt that the 0.5 reaction type is superior for compact design. Its advantages are particularly great at the lower stages, where sound velocity is low and the better flow coefficients allow the large volumes to be handled with ease. If the compressor has many stages and a high compression ratio, conditions change at the higher stages. The Mach-number limitations, Equation [32], are critical only for low sound speeds. The maximum permissible pressure rise per stage increases obviously with the square of the velocity of sound, as shown by Equations [32] and [33], and consequently, with the ratio of the absolute temperature increase. After a certain number of stages, therefore, the temperature rise is sufficient to annul the advantage of the 0.5 reaction type, and from then on, spinless rotor-outlet bladings can be used without fear of compressibility trouble. Owing to the greater τ_{d0} of this type, the pressure rise per stage is then actually better than with a 0.5 reaction blading. Besides, owing to the decrease in specific volumes, the axial velocity can be diminished conveniently; and, as already seen, this fits the spinless outlet type very nicely.

For compressors having a high compression ratio and using free-vortex bladings, the combination of 0.5 reaction bladings at the lower stages and of spinless outlet bladings at the higher stages is particularly indicated. The flow-smoothing qualities of the latter, due to acceleration in the stator, come particularly handy at the later stages, where the flow tends to get very unruly indeed. This combination allows a greater number of stages to be fitted without intercooling; gives a higher mean-pressure rise per stage, higher stage efficiencies; and the leaving losses from the last stage are as low as possible.

There is no doubt that this type of compressor is the most efficient and compact. The only snag might appear to be the diversity of blading. This is only an apparent complication, however, because in free-vortex high-efficiency compressors the blading cannot be standardized in any case; it must differ constructively from stage to stage, or at least from small group to small group of stages. Consequently, standardizing the blading diagrams is only a way of pandering to drawing-office laziness, but represents no real advantage in production.

Constant-Reaction Blading

Up to now, only the continental types of blading, illustrated in Figs. 1 to 3, have been examined. In the comparison of 0.5 reaction and spinless outlet types, in the preceding section, it has been seen, under (*a*) and (*b*), that the free-vortex blade twist causes changes in the degree of reaction which affect unfavorably the 0.5 reaction type. If the degree of reaction remained 0.5 from tip to hub, then conditions would be improved, and both a higher u_0 and a higher τ_{d0} could be used. Of course the blade still would have to be twisted, because the deviation and τ_d must increase from tip to hub, to insure an even pressure rise at all diameters, just as in the case of continental-type blades. Further, if axial velocity remains constant at all diameters, φ of course will increase from tip to hub. But this type of blading does not respect the other continental condition; that free-vortex motion should prevail in the axial gaps between stator and rotor. In other words, this is not a free-vortex blading, because conditions in the gaps do not comply with the free-vortex law ($c_u \times d$ = const for all diameters). Consequently, this type

of blading is not included in continental types. According to theory, it should give low efficiencies. In practice, however, this type of blading is widely used in England, where it originated, with excellent results. It is called "constant-reaction" type in England, see Howell (1), and "symmetrical stage" here, see Ponomareff (6).

When a very compact compressor with very few stages is desired, this type of blading gives excellent results, and the efficiency is much higher than continental theory would expect. Of course the shorter the blades, the smaller the difference between conditions prevailing with 0.5 reaction free-vortex blades and constant-reaction blading. For long blades, the British have developed the so-called "half-reaction" type. This is a compromise between constant reaction and free vortex 0.5 reaction; R varies on getting farther away from the pitch diameter, but the variation is about one half that required by free-vortex conditions. A good description can be found in Howell (1); this type too has proved quite satisfactory in practice.

In the United States, another and more unorthodox type of blading has been developed recently, namely, the solid rotation blading. In it, the tangential component of the absolute velocity increases proportionately with the radius (instead of being inversely proportional to the radius, as in the free-vortex type). The fluid rotates solidly with the rotor, hence the name. Of course in this case radial equilibrium is possible only if the axial velocity is not constant at all radii. By suitable variations of axial velocity, radial equilibrium can be obtained; extensive calculations in three-dimensional potential theory are required. The solid rotation type violates all canons of continental theory; but apparently it gives excellent results, allowing high pressure rises per stage, a very large flow per unit cross-sectional area, and relatively cheap and easily built blades, owing to the limited twist, even when the blades are very long. It is of course particularly suitable for aviation work. It would be very nice if some of the American proponents of this type of blading were to write a paper about it; the published material available at present is scanty indeed.

Continental Theory Discussed

The fact that constant-reaction, half-reaction, and solid rotation bladings can give quite reasonably high efficiency shows that the premises of continental theory are not necessarily infallible. Consequently, a critical examination is desirable. It has been seen under the heading, "Conditions Along the Radius," that the main premises are: (*a*) constant pressure rise in the stage, so as to avoid radial pressure differences along the radius (this is obtained by constant circulation along the blade); (*b*) radial equilibrium in the gaps (this is obtained by free vortex conditions); (*c*) the general assumption that axial velocity is constant along the radius.

It must be noted that free-vortex condition (*b*) includes constant circulation (*a*), because if $(c \times d) = \text{const}$ along the radius, then, obviously, the constant-circulation condition $(c_{u1} - c_{u2})t = \text{const}$ will be satisfied, as the pitch t is proportional to d; but the obverse is not true because $(c_{u1} - c_{u2})\ t = \text{const}$ does not imply any hypothesis on the values of c_{u1} and c_{u2}.

In constant-reaction and similar blading, condition (*a*) is more or less satisfied, while condition (*b*) generally is not. Practical results prove that the consequences are not so important as regards efficiency. Undoubtedly, this is due in part to the fact that axial gaps between stator and rotor at the present time are much smaller than when the continental theory was formulated by Keller in 1934. Also, radial equilibrium can be re-established by corrections in axial velocity along the radius, as noted previously.

Also, there is another point where continental theory is weak, however; and in it lies the explanation of the good results given by constant-reaction blading notwithstanding its theoretical failing. The point is this: Condition (*a*) given is very important; the fact that pressure should rise steadily across the whole section prevents radial pressure differences which would cause longitudinal whirls and eddies and consequent loss of efficiency. (It must be stated clearly that by pressure is meant static pressure and not total energy pressure, as the latter cannot by its nature cause radial disturbances). Now, though this condition (*a*) is set as a basis of continental theory, actually, it is satisfied at one point only (the end of the stage after the rotor) in nearly all types of continental blading. This is due to the fact that the degree of reaction varies along the blade; and that is why so much importance has been given to R in this paper. In spinless rotor-outlet bladings, the variation of R are small, for the reasons given under "Blading Types Compared." Fig. 2 (*a*, *b*, *c*) shows that the degree of reaction in $R_s = 1.064$ at the tip, $R_0 = 1.1$ at the pitch, $R_N = 1.177$ at the hub; this means that, in the stator, pressure decreases more rapidly at the hub than at the tip and is then built up again more rapidly at the hub in the rotor. The differences, however, are small; and they are also small in the spinless stator-inlet type, Fig. 3(*a*), while there is no difference at all in the symmetrical stator type, Fig. 3(*b*), where $R = 1$ all along the radius. The fact that Keller, the father of modern continental theory, worked mainly on spinless types of blading probably is the cause of these pressure differences being neglected (just as the differences caused by efficiency variations along the radius had been neglected without appreciable error).

However, when we come to consider the 0.5 reaction-type free-vortex blading, things become quite different. As shown in Fig. 1, the degree of reaction varies notably; from $R_s = 0.68$ at the tip, it becomes $R_0 = 0.5$ at the pitch, and only $R_N = 0.115$ at the hub. This means that in the stator the pressure rise is small at the tip and high at the hub (i.e., contrary to centrifugal-force conditions) and vice versa in the rotor; and the pressure differences between tip and hub are very considerable indeed. The only point where pressures are the same in all the sections is at the end of the stage. There is no doubt that these differences of pressure must cause disturbances and even eddies, the more so, the smaller the number of stages used to attain a certain compression ratio, because the pressure differences are proportional to the stage pressure rise. It also must be noted that these pressure differences are highest at the axial gap between stator outlet and rotor inlet. Even though free-vortex conditions prevail there, radial equilibrium is not satisfied, because free-vortex conditions give radial equilibrium only when static pressure is the same at all radii. This is true only in the gap between rotor and stator.

On the other hand, constant-reaction bladings of course have no radial pressure differences, because the degree of reaction is the same at all diameters. In the author's opinion that is why constant-reaction bladings give better results than the theory expects; what is lost in the rotor-outlet gap (because radial equilibrium is not observed there), is gained because of the steady pressure rise at all diameters and along the whole stage.

Conclusion

From the foregoing considerations, two practical conclusions can be drawn:

(*a*) In the case of a compressor with a high compression ratio (6 to 1 or more) and without intercooler, the best results will be obtained by the use of three different types of blades, i.e., constant-reaction (or solid rotation) for the low-pressure section, 0.5 reaction free vortex for the medium-pressure section, and spinless rotor-outlet free vortex for the high-pressure section, each type being particularly well-suited for its work. This ar-

rangement is particularly useful in high-altitude aviation work, where conditions are very critical to Mach number owing to low sound velocity.

(*b*) It would be very interesting if comparative tests were made between these three types of blading. These tests could be carried out at low Mach numbers, because what is wanted is a measure of the radial equilibrium and pressure distribution along the stages in the three types, and of the smoothing properties of the spinless rotor-outlet type (these could be tested by causing artificial disturbances before the stator inlet); and a measure of the real degrees of reaction and of the radial distributions of axial velocity. These tests would constitute an interesting and very useful bit of work for some university gas-turbine laboratory.

Centrifugal Compressor. The author desires to make it clear that although this long paper deals only with axial compressors, he does not believe that the axial compressor is the only type suitable for high-efficiency gas turbines. On the contrary, he is convinced that if properly developed, the centrifugal compressor or a combination of one or more axial stages, followed by one centrifugal stage will probably give very satisfactory results.

Acknowledgment

The author wishes to express his deepest gratitude to Prof. H. Quiby of the Federal Polytechnic School in Zurich, whose lectures on gas turbines and compressors are the best on the Continent, and whose unexcelled experience and unfailing courtesy are always of the greatest help to all who ask his advice.

BIBLIOGRAPHY

1 "Design of Axial Compressors," by A. R. Howell, in "Lectures on the Development of the Internal Combustion Turbine," Proceedings of The Institution of Mechanical Engineers, 1945; ASME, reprint, 1947.

2 "The Theory and Performance of Axial Flow Fans," by C. Keller and L. S. Marks, McGraw-Hill Book Co., Inc., New York, N. Y., 1937.

3 "Neue Allgemeine Theorie der Mehrstufigen Axialen Turbomaschinen," by Walter Traupel, Leemann & Co., Zurich, Switzerland, 1942.

4 "A Marine Gas Turbine Plant," by C. R. Soderberg and R. B. Smith, Proceedings of The Society of Naval Architects and Marine Engineers, vol. 53, 1945, pp. 249–268.

5 "Principles of Jet Propulsion," by M. J. Zucrow, John Wiley and Sons, New York, N. Y., 1948.

6 "Axial-Flow Compressors for Gas Turbines," by A. I. Ponomareff, Trans. ASME, 1948, vol. 70, pp. 295–304.

7 "Le Turbine a Gas," Note estratte dal corso tenuto a Roma al Consiglio Nazionale delle Ricerche nel Maggio-Luglio 1947, dal Dr. Ing. Franco Martinuzzi a cura dei Capitani dil Genio Navale Ingg. Bartorelli, Chiappi e Toddini; published in "Note di Ingegneria Navale," pubblicazionne riservata della Marina, Roma; in preparation.

Discussion

R. S. Hall.[5] It does not seem at all surprising that the author has found little evidence to indicate either European influence on the design methods of American manufacturers or mutual influence among these same American manufacturers at the present time. Most of the American manufacturers active in the gas-turbine field laid the foundations for their design methods either prior to or during the early part of the war, at which time there was practically no available information on design practice. For example, most American axial-flow-compressor designers began with Curt Keller's book,[6] stirred in a large portion of their own basic thinking, added a small portion of research testing, and came up with a design method. Although considerably more information is now available, most of the American design approaches were fairly firmly established before it became available. Because of wartime restrictions, there is even today little exchange of basic information between the American manufacturers.

In commenting on the more detailed treatment of axial-flow-compressor calculation methods the writer would like to make clear the natural differences in design philosophy between the stationary gas turbine and the gas turbine for aircraft propulsion. The three basic factors, size, weight, and efficiency, determine the proper "design philosophy" for a particular type of application.

For the stationary gas turbine, the first two factors exert a relatively minor influence, the efficiency being by far the most important in determining the "design philosophy." This relative weighing leads to two fundamental distinguishing features in the design approach for the stationary plant. (*a*) That the efficiency of the compressor must be maintained at as high a level as possible even at some expense of size and weight. (*b*) The research effort should be confined to obtaining the last possible bit of efficiency from the most fundamentally sound approach.

This leads naturally to a compressor design characterized by: (*a*) constant axial velocity at all radii; (*b*) "free-vortex" type of tangential velocity distribution between blade rows.

This type of design represents the closest practical approach to the ideal "potential" flow, and therefore the most fundamentally sound approach to high efficiency. The author's suggested design variation employing "spinless rotor outlet," shown in his Fig. 2, represents an effort to obtain the "last possible bit of efficiency from the most fundamentally sound approach."

In the case of the aircraft gas turbine, the picture is considerably different. We must go much deeper into the scheme of things in order to obtain the smallest possible size and weight consistent with acceptable efficiency levels. While we should not cast aside arbitrarily the fundamentally sound approach, we must investigate thoroughly the possibilities of deviations from this approach to ascertain the magnitude of losses introduced and finally achieve a fine balance between size, weight, and efficiency.

These considerations lead us to a compressor design characterized by: (*a*) axial velocity which varies with radius; (*b*) "forced-vortex" type of tangential velocity distribution between blade rows.

This type of design introduces a host of additional unknown factors which will take a considerable amount of time and money to evaluate. This evaluation, undoubtedly, will be done almost exclusively by the aircraft gas-turbine builders.

These two basic classifications seem to exclude one commonly mentioned type of compressor design; the "symmetrical stage" with constant axial velocity. This exclusion is a natural one for such a design does not satisfy the condition of equilibrium between the centrifugal force and the radial pressure gradient, and, while diagrams may be established on such a basis, it seems doubtful whether in fact a compressor actually can operate in this manner.

This paper seems to be a very timely summary of continental compressor-design practice. It does seem to be representative of stationary gas-turbine practice primarily.

If a survey could be conducted, in which the distinction between stationary and aircraft practice were carefully made, it is quite probable that we would find minor differences in routine calculation procedures but a fairly good general agreement on basic principles, and fundamental limitations and their application, among those active in each of these two gas-turbine fields.

[5] Aircraft Gas Turbine Engineering Division, General Electric Company, Schenectady, N. Y.

[6] Author's Bibliography, ref. (2).

J. T. Rettaliata.[7] The author has made a worth-while contribution to the literature on axial-flow gas turbines and compressors. The limited amount of published design information on this subject assures a cordial reception to a paper of this kind.

The use of certain simplifying assumptions for the purpose of facilitating the calculation procedure is adequately explained. In particular, the assumption of constant density, or incompressibility, is valid when confined to a single stage where the change in energy is relatively small. Where additional stages are involved, a series of appropriate adjustments in the state properties of the working fluid should be made. By imposing the condition of incompressibility, simplified dimensionless quantities are obtainable which permit a ready prosecution of the design.

Specifically, it is mentioned in the paper that the pressure coefficient ψ, a dimensionless quantity, may be expressed as

$$\psi = \frac{\Delta p}{\frac{1}{2}\rho u^2}$$

where

Δp = stage pressure rise
ρ = mass density
u = mean blade velocity

Although not stated in the paper, another dimensionless quantity is the temperature-rise coefficient, which in this discussion will be referred to as ϕ_t, and it can be shown to be equivalent to

$$\phi_t = \frac{J\, c_p\, \Delta t}{\frac{u^2}{2g}}$$

where

Δt = stage temperature rise
c_p = specific heat
u = mean blade velocity
g = gravitational constant
J = Joule constant

By suitable operations the very useful relationship of

$$\psi = \eta\, \phi_t$$

may be obtained, where η is the stage efficiency. It will be noted that the numerator of the right-hand member of the temperature-rise-coefficient expression is the work input to the stage.

As mentioned in the paper, by assuming the blade rotational velocity to be unity at all radii, it is possible to depict the degree of reaction as a dimension on the velocity triangles. The author is to be congratulated on revealing this useful concept as well as other clever manipulations contained in the paper.

[7] Dean of Engineering, Illinois Institute of Technology, Chicago, Ill. Mem. ASME.

Author's Closure

There is no question that the approach to aircraft gas-turbine design is radically different from that of industrial gas-turbine design, for the reason so clearly exposed by Mr. Hall. In the author's opinion it seems, however, probable that this difference in outlook will gradually disappear in the future, particularly if the aircraft industry will require units of very high power. Alterations in aircraft structure, such as putting the engines inside the main fuselage, would probably bring the aircraft gas turbine nearer to the high-efficiency gas turbine by reducing the need of a small outside diameter.

The high efficiency obtained in continental industrial gas turbines is undoubtedly possible for another reason also—the large gas volumes and the consequently long compressor and turbine blades employed. The author has recently (October–December, 1948) noted this while visiting the latest Swiss gas-turbine installations (the 13,000-kw and 27,000-kw Brown-Boveri turbines at Beznau are particularly good examples). The blades are so long that the influence of clearances and of rotor and stator boundary layers on total efficiency are very much smaller than would be the case in machines dealing with a normal flow of air. The author hopes that Mr. Hall and other exponents of modern American aircraft gas-turbine practice will publish papers about the type of compressor mentioned by Mr. Hall, that is, one having (*a*) axial velocity which varies with radius and (*b*) "forced-vortex" type of tangential velocity distribution between blade rows. The amount of information that can be obtained at present on these subjects is small, and further papers, particularly by people who, like Mr. Hall, have both extensive practical experience and high mathematical abilities, would be extremely welcome.

Dr. Rettaliata has rendered a useful service by reminding readers that actual compressor design has thermodynamic aspects which are equally as important as the fluid-mechanic considerations. In actual calculations, thermodynamic coefficients such as ϕ_1 mentioned by Dr. Rettaliata are extremely useful. A good way of calculating compressor design is to first make the calculations in fluid mechanical terms, and then to recalculate it in thermodynamic terms as a check. The author has not discussed the thermodynamic aspects in his paper simply to avoid introducing more symbols and concepts in a paper which is necessarily very condensed.

In closing, the author wishes to express his appreciation to Mr. Hall and Dr. Rettaliata for their contributions to the worth of his paper.

Tests on V-Belt Drives and Flat-Belt Crowning

By C. A. NORMAN,[1] COLUMBUS, OHIO

The investigations here reported were undertaken with the liberal support of the Goodyear Tire and Rubber Company, at the instigation of the company's well-known belt engineer, the late Mr. Ray S. Carter. The immediate reason for the investigation was the desire to clear up under just what tension the front-end belt drives on certain important automotive vehicles ought to be applied to insure dependable power transmission even under sustained heavy load, and yet insure a reasonable belt life. These drives involve high speeds, small contact angles, and also small pulley diameters. Hence it was necessary to clear up the influence of contact angle on transmissive power with the small pulley diameters used. Also belt-life determinations were in order. The project then involved, in fact, all the fundamentals of power transmission with modern V-belts. While the work was in progress Mr. Paul D. Suloff, who throughout maintained current liaison between the company and the experimenters, suggested that the effectiveness of pulley crown on modern flat belts should be investigated, since with these belts customary crown practice might not be salutary for the belt. Such investigations were then added to the program. In September, 1945, the project suffered a very serious loss in the death of Mr. Carter at the height of his powers and usefulness. The work, however, continued under the detail supervision of Mr. Suloff as before, and by the spring of 1946, it was felt that the fundamentals had been determined; but for one belt size only. Experiments on other belts continued into 1947, at which time it was necessary for the author to curtail his activities. This made it impossible to give a complete experimental answer to the practical question originally posed by Mr. Carter. However, so much fundamentally important material had been accumulated that a delay in making it available to belt users would not seem justifiable. Supplementary tests **(see p. 347)** **extend the findings to higher loads and speeds.**

Details of Tests

Centrifugal Force

THE basic formula used in this country for belt transmission has been

$$\frac{T_1 - \frac{wv^2}{g}}{T_2 - \frac{wv^2}{g}} = e^{f\alpha/\sin\gamma} \quad \ldots\ldots [1]$$

where T_1 = tight-side tension, lb
T_2 = slack-side tension, lb
w = weight of belt, pounds per foot of length
v = belt speed, fps
g = gravity acceleration, fps^2
e = basis of natural logarithms
α = contact angle, radians
f = coefficient of friction
γ = one half the groove apex angle; $\gamma = 90$ deg for flat pulleys

The formula is derived on the assumption that while the belt runs over the pulley, both tensions are reduced by the amount wv^2/g, and that there is nothing to compensate for this reduction. According to this formula, if wv^2/g equals the slack-side tension T_2, power transmission ceases, since $T_1 - wv^2/g$ would then also have to become zero, $e^{f\alpha/\sin\gamma}$ not being equal to infinity.

This, as is now becoming realized in this country, and was realized earlier in Europe, is wrong in fundamental theory and amounts in practice to a mere approximation which is more and more in error the slacker the belt can run, and the less it stretches elastically.

The point is this: the centrifugal tension term wv^2/g contains neither the radius of curvature of the belt, nor the contact angle. Hence it is independent of both. Now the belt hangs between the pulleys to some sort of a curve, which for a perfectly flexible belt and a horizontal or moderately inclined drive would be a catenary. Hence there is also in this part of the belt a centrifugal tension wv^2/g which adds itself to the tension resulting from the weight. This tension increment between the pulleys would compensate exactly for the centrifugal reduction on the pulleys; and, if the belt did not stretch, the centrifugal force would have no influence upon the transmissible power at all.

The centrifugal-force tension however, stretches the belt, and while this has no effect on the centrifugal tension itself, it does have an influence on the tension caused by the weight of the belt. The tension will decrease with increased sag, i.e., with the stretching of the belt. Hence since the centrifugal tension stretches the belt, if it is at all stretchable, it will reduce the tension due to weight. Therefore formulas to allow for influence of centrifugal force must contain terms reflecting the influence of the tension on sag and belt length.

The complete formula for this influence for a horizontal idling belt was given by the author in a previous paper.[2] The derivation is lengthy and will not be reproduced here. The formula, however, has the following form

$$T_0 - T_s = \frac{wv^2}{g} - \frac{E}{l_0}(L_s - L_0) - \frac{w^2 L_0^3 E}{24 l_0}\left(\frac{1}{T_s^2} - \frac{1}{T_0^2}\right) - \frac{0.132K^{3/2}}{r^2}\frac{E}{l_0}\left(\frac{1}{T_s^{3/2}} - \frac{1}{T_0^{3/2}}\right) \quad \ldots\ldots [2]$$

in which

T_0 = initially applied standstill tension
T_s = actual shaft pull per side with belt running. If total tension per side is T, we have $T_s = T - wv^2/g$
E = elasticity constant; equal to force which would stretch belt elastically 1 in. per in.

---

[1] Professor of Machine Design, Department of Mechanical Engineering, The Ohio State University. Mem. ASME.

Contributed by the Machine Design Division and presented at the Annual Meeting, New York, N. Y., November 29–December 3, 1948, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—A-102.

[2] "High-Speed Belt Drives," Bulletin No. 83 by the Engineering Experiment Station, Ohio State University, 1934.

l_0 = original unstretched length of one half of belt
L_s = shaft distance with belt running
L_0 = shaft distance with belt stationary
K = flexibility constant, such that bending moment needed to flex the belt to a radius of curvature r is K/r

The difference between the standstill shaft pull per side T_0, and the running shaft pull T_s is equal to the old-time quantity wv^2/g, modified, however, by three terms.

One term allows for a difference in shaft distance. Naturally, if the shaft distance decreases, the sag increases and reduces the shaft pull, entirely apart from the influence of centrifugal force. In a floating drive with special tensioning device this may well occur as will be shown later.

Another modifying term allows for elastic stretch in the belt. The greater E is, i.e., the less stretchable the belt is elastically, the greater is the modifying effect of this term, the smaller the difference between T_0 and T_s. The same effect is produced by great belt weight w per inch of length, and by long center distance. This is all on the assumption that T_s is smaller than T_0, i.e., that the term wv^2/g predominates.

The final modifying term allows for belt stiffness. If K is great, i.e., if the belt is stiff, then the radius of curvature of the belt at the pulley is great, and for a given belt length, the belt length on which sag may occur is reduced. Therefore stiffness decreases the slackening effect of centrifugal force.

It should be clear that Equation [2] is not meant to be used for computing purposes. The derivation is approximate, applies to an idling belt only, and the constants E and K are usually not known to the belt user. Neither is w. The formula is meant only to make clear in a qualitative, or but roughly quantitative, way the relations that obtain. Actual data on the influence of centrifugal force must be obtained by direct tests. In the tests previously reported,[2] the belts and pulleys as well as the speeds used were such that only the influence of the term $E(L_s - L_o)/l_o$, allowing for changes in shaft distance, was of well-measurable magnitude. The findings were in fair agreement with computed expectations. The influence of the other terms was shown in tests here to be reported.

Before proceeding to this we shall, however, remark on one point, namely, that the centrifugal force as such does not bulge the belt and, in fact, gives the belt between the pulleys no definite shape whatever. A chain mounted on vertical shafts takes the shape shown in Fig. 1. The bending-in where the chain leaves the pulley is due to inertia effects in the links, not to centrifugal force. It can be eliminated by arranging the pivot points of the links so that the chain tension exerts a righting moment.

A chain hanging slack from a sheave on a horizontal shaft will not bulge but can be thrown into all manner of shapes by light knocks or pressure.

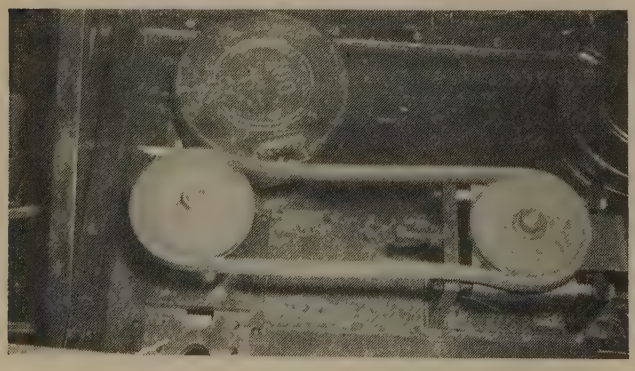


Fig. 1 Chain Mounted on Vertical Shafts

A perfectly flexible belt would behave similarly. If a belt bulges, this is due merely to stiffness, and actual tests show that the bulging force is no greater at 7000 fpm than it is at standstill.

All this is of course in harmony with the fact that neither the radius of curvature nor the included angle appears in the formula for the centrifugal tension wv^2/g.

Tests. In Fig. 2 is shown a plot of idling tests run to verify the stretching action of the centrifugal force. The driven pulley is mounted on a sliding carriage. The pulleys were of 9 in. OD. The belt was a C-105 smooth cotton V-belt, which is fairly

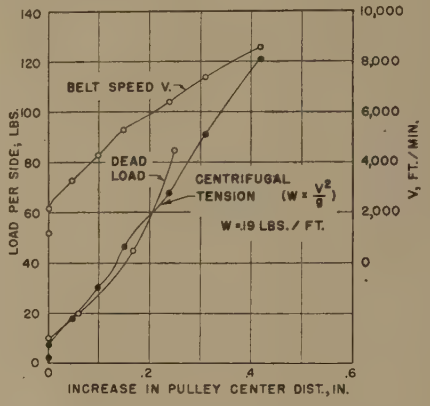


Fig. 2 Centrifugal-Force Tests With Two Pulleys, Using Goodyear C-105 Smooth Cotton V-Belt
(9-in. average pulley OD; dead-load stretch test made at 70 rpm, and driven pulley on sliding carriage.)

stretchable elastically as modern belts go. The pull on the driven shaft was first held barely sufficient to keep the belt running straight. The increase in center distance as the belt was speeded up to a speed of over 8000 fpm is plotted as abscissa. The belt was then idled slowly at 70 rpm, and dead loads were applied at the driven end to produce a stretch equal to that produced by the previous centrifugal force. The practically perfect agreement between the computed centrifugal tension wv^2/g and the dead load per side for equal stretch is noted.

The really important test results are plotted in Figs. 3 to 8, inclusive. Figs. 3 to 5, inclusive, show that for two different lengths of steel cable V-belts, i.e., 96 in. and 144 in., respectively, at two different initial tensions, it is possible to transmit not only as much, but actually more effective pull at high speed than at low speed. This applies up to speeds of over 6000 fpm. The slips, while they increase with the load, do not seem to increase with the speed for constant load, at least not for the 96-in. belt. Now steel-cable belts stretch very little elastically. However, as Fig. 6 shows, the same situation obtains for a 144-in. long C-size cotton-strand V-belt. The absolute values of the slip are greater than for the steel belt, but do not increase with the speed for equal effective pulls.

In addition, Figs. 3 to 6, inclusive, reveal the startling fact that the actual effective pull, $T_1 - T_2$, transmissible exceeds the initial tension $T_1 + T_2$ at which the shaft distance was locked. About this more later.

The tests just discussed were carried out on horizontal drives with the driven shaft journaled in a sliding carriage, and the torque taken through a flexible shaft to a Prony brake. The shaft distance was locked after the initial tension had been set by dead weights.

In the tests reported in Figs. 7 and 8 the driven shaft carried a calibrated fan. The pulleys were the same 9-in-OD pulleys as before. The belts were still C-size, but only 68 in. long. They were of steel-cable type, as well as cotton-cord type, and of smooth

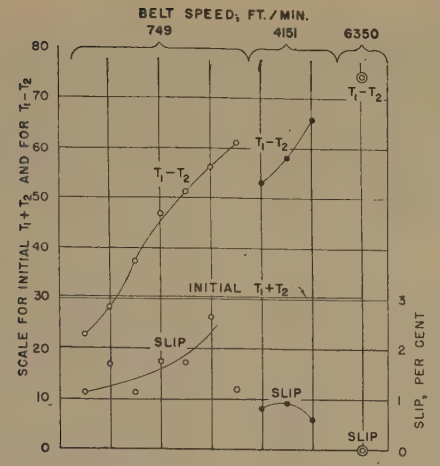


FIG. 3 COMPARATIVE SPEED TESTS USING PRONY-BRAKE LOAD WITH GOODYEAR C-96 SMOOTH STEEL-CABLE V-BELT
(Locked center distance; 9 in. average pulley OD; and two-pulley drive direct from dynamometer; nominal contact angle 180 deg.)

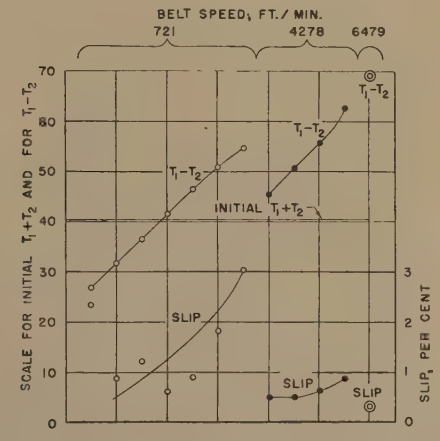


FIG. 4 COMPARATIVE SPEED TESTS, USING PRONY-BRAKE LOAD, WITH GOODYEAR C-96 SMOOTH STEEL-CABLE V-BELT
(Locked center distance; 9 in. average pulley OD, and two-pulley drive direct from dynamometer; nominal contact angle 180 deg.)

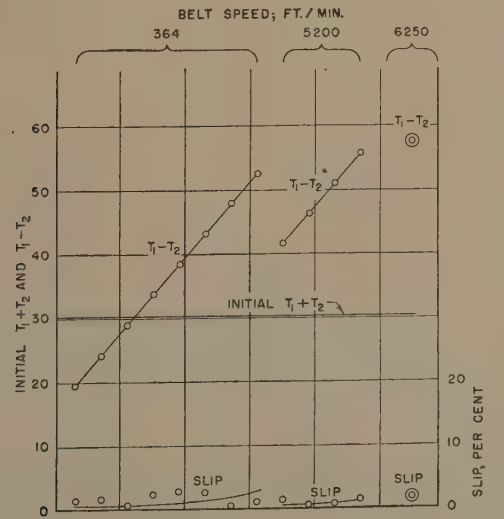


FIG. 5 COMPARATIVE SPEED TESTS, USING PRONY-BRAKE LOAD, WITH GOODYEAR C-144 SMOOTH STEEL-CABLE V-BELT
(Locked center distance; 9 in. average pulley OD, and two-pulley drive direct from dynamometer; nominal contact angle 180 deg.)

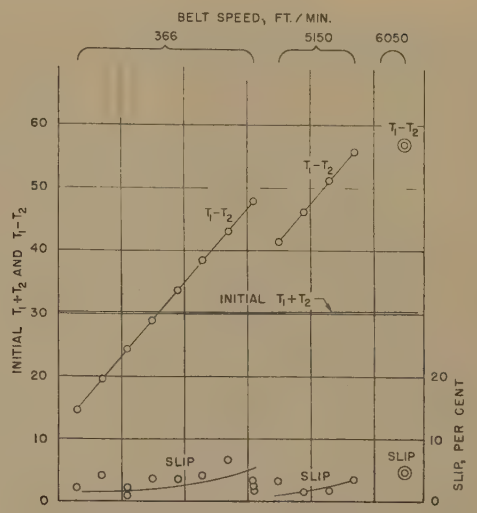


FIG. 6 COMPARATIVE SPEED TESTS, USING PRONY-BRAKE LOAD, WITH GOODYEAR C-144 SMOOTH COTTON V-BELT
(Locked center distance; 9 in. average pulley OD, and two-pulley drive direct from dynamometer; nominal contact angle 180 deg.)

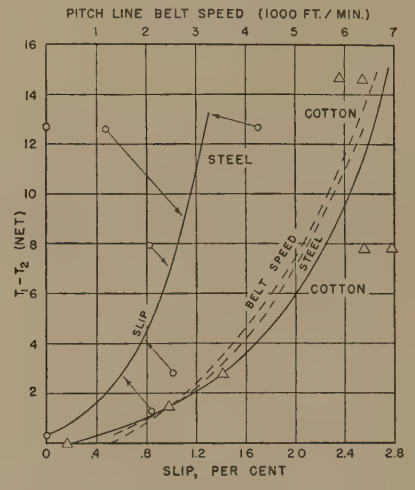


FIG. 7 COMPARATIVE HIGH-SPEED TESTS, WITH 9-IN-OD PULLEY; $8^1/_4$ IN. PITCH DIAM ASSUMED; C-68 NOTCH BELTS, AND DRIVE LOCKED AT $T_1 + T_2 = 30$ LB
(T_1 — T_2 effective pull consumed by fan only.)

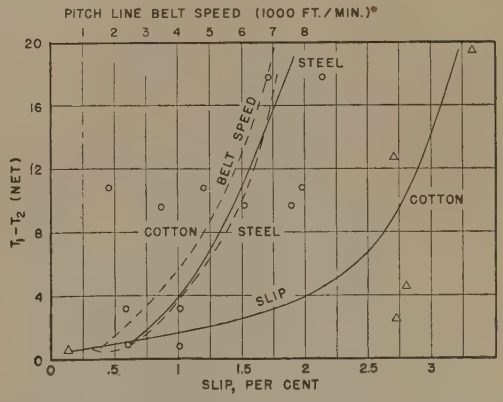


FIG. 8 COMPARATIVE HIGH-SPEED TESTS, WITH 9-IN-OD PULLEY; FAN LOAD $8^1/_4$ IN. PITCH DIAM ASSUMED; SMOOTH C-68 BELTS, AND DRIVE LOCKED AT $T_1 + T_2 = 30$ LB
(T_1 — T_2 effective pull consumed by fan only.)

V-type as well as notch type. The main interest centers in showing the very much greater slip of the cotton-cord belts. The slip was also greater for the stiffer smooth belts than for the more flexible notch belts.

The fact that a belt on a fixed shaft distance can transmit more effective pull at high speed than at low is of course no new discovery. The theoretical basis for this situation was at least in part given by Friedmann[3] as far back as 1894.

Kammerer[4] proved it by tests at very high speeds the results of which were published in 1907.

These tests have been referred to by the author in an earlier publication.[5] In 1926 G. Schulze-Pillot published a booklet[6] giving a quite complete, although cumbersome, theory for the influence of belt stretch and centrifugal force.

This book is mentioned in an earlier reference[2] in this paper and additions to the theory were given as already indicated.

The matter has been referred to in some recent textbooks on machine design and also in other publications. Yet the author does not believe that it has been emphasized sufficiently that, in Europe, the matter has been so well established by practical experience that a very leading belt firm, G. Otto Gehrckens in Hamburg, actually recommended effective pulls on sufficiently large pulleys increasing all the way up to a belt speed of 10,000 fpm.[7]

Comment. Since the complete mathematical theory, as given previously[2] may seem somewhat complicated, it may be well to point out that the centrifugal tension in the belt between the pulleys is of course demanded simply by Newton's third law, stating that an action is balanced by an equal reaction. Assume, for simplicity, a horizontal drive with equal pulley diameters. Let the applied tension T in the belt be zero. The outward directed centrifugal force at each end sets up a tension wv^2/g which tends to pull the belt apart. As long as the force does not rupture the belt, this tension must be balanced by an equal force wv^2/g in the belt between the pulleys.

If the belt does not stretch, the intimacy of contact between the belt and the pulley is not changed by the centrifugal tension. The pressure between pulley and belt will be determined by any applied initial tension T_0. If the belt stretches, the tension T_0 will be influenced by the sag, and the belt pull against the pulleys will be determined by the sag condition.

With a vertical belt, the stretch due to centrifugal tension actually may pull the belt free of the lower pulley. A well-known illustration of this fact was given by Goodman[8] many years ago. His experiment can be easily duplicated by a rubber V-belt having no cords, and such belts were run in our tests.

Fixed and Floating Shaft Distance

The new and somewhat startling thing shown in Figs. 4 to 6, inclusive, as already pointed out, is that with fixed shaft distance, it is possible to transmit an effective pull $T_1 - T_2$ greater than the initially applied total tension $T_1 + T_2$. This had been observed also by Goodyear engineers at Akron. Since the pressure of war work did not permit a comprehensive investigation of the phenomenon at Akron, the tests at The Ohio State University were suggested by Mr. Carter.

One possible explanation of this phenomenon which, however, will apply only to certain types of belts, will immediately suggest itself. In a paper,[9] published in 1909, Carl G. Barth suggested a formula for short or vertical belt drives to cover the relation between the initially applied tension per side T_0, and the working tensions T_1 and T_2

$$\sqrt{T_1} + \sqrt{T_2} = 2\sqrt{T_0} \quad [3]$$

However, this formula applied only to belts, the stress-strain curve of which would run as curve A in Fig. 9, that is to say, the relation between belt length l and tension T could be represented by an equation $l = a + b\,T^m$, where a and b are constants, and m is an exponent less than unity.

For such a relation, as Fig. 9 shows, an increase in tension from T_0 to T_1 on the tight side would correspond to a smaller stretch than for an equal amount of tension drop from T_0 to T_2 on the slack side. Hence as load was applied, the belt would shorten. If the shaft distance were kept constant, the belt would tighten.

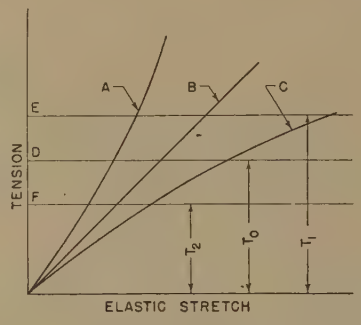


Fig. 9 Stretch-Strain Curves

If the stress-strain relation is linear, as represented by curve B, then the extension on the tight side would equal the shortening on the slack side, and the belt length would remain constant. We should have

$$T_1 + T_2 = 2\,T_0 \quad [4]$$

Finally, if the stress-strain curve bends toward the strain axis, as curve C in Fig. 9, the extension on the tight side would be greater than the contraction on the slack side. The belt length would increase with the load and we should have

$$T_1^n + T_2^n = 2T_0^n \quad [5]$$

where n is greater than unity.

The total tension under load would be smaller than the initial tension, and the effective pull transmissible with fixed shaft distance would be smaller than the load transmissible with a floating drive.

In stress-strain determinations for V-belts at the Ohio State University, it was found that the majority of them had stress-strain curves of type C. In spite of this they transmitted more load with fixed shaft distance than with floating drive.

The matter was clarified by arranging to have the driven shaft mounted on a sliding carrier as in Fig. 10. The driven shaft either carried a fan or was connected through a flexible shaft to a Prony brake. Tests were run with C-105 V-belts on 9-in. pulleys, and with C-68 V-belts on 6.9-in-OD pulleys. Tests developed the following results:

[3] "Beitrag zur Berechnung des Drahtseiltriebes," by C. Friedmann, *Zeitschrift des Vereines deutscher Ingenieure*, vol. 38, July 28, 1894, p. 891.

[4] "Versuche mit riemen- und Seiltrieben," by W. Kammerer, *Zeitschrift des Vereines deutscher Ingenieure*, vol. 51, July 13, 1907, p. 1085.

[5] "Principles of Machine Design," by C. A. Norman, The Macmillan Company, New York, N. Y., 1925, p. 410.

[6] "Neue Riementheorie," by G. Schulze-Pillot, Julius Springer, Berlin, Germany, 1926.

[7] Reference 5, p. 410.

[8] "Mechanics Applied to Engineering," by J. Goodman, Longmans, Green & Company, New York, N. Y., 1904 edition, Fig. 269.

[9] "The Transmission of Power by Leather Belting," by Carl G. Barth, Trans. ASME, vol. 31, 1909, p. 29.

When the tension on the slack side approaches zero, then, if excessive slip does not occur, the drive becomes a hoist, the tight side pulling the driven pulley toward the driving pulley. This of course increases the slack and decreases the contact angle on the pulleys so that increased slip and finally breakdown of the drive occur. To keep the shaft distance constant, it is necessary to increase the load L. This of course increases the contact angle, as well as the tensions, and permits the transmission of a greater effective pull.

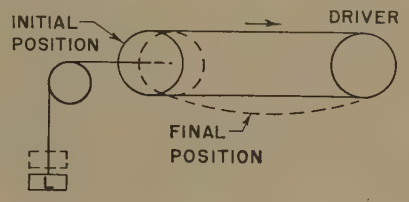


Fig. 10 Floating Belt Drive

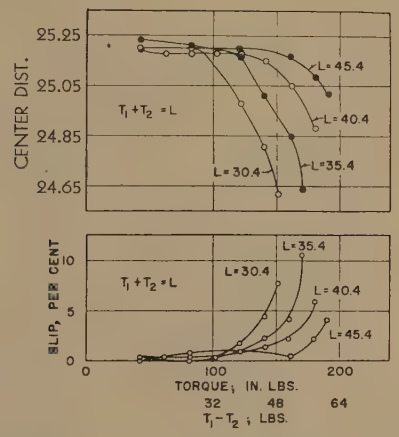


Fig. 11 Change in Center Distance With Sliding Pulley, Using Prony-Brake Load, Goodyear C-68 Smooth Steel-Cable V-Belt

(6.9 in. average pulley OD; 450 rpm approx.)

Hence a fixed shaft distance is equivalent to adding to the total tension and permits the transmission of a greater effective pull without excessive slip than does a floating drive.

This effect is shown diagrammatically in Fig. 11 which refers to a horizontal floating drive. The shaft distance and the slip are plotted against the torque transmitted, or against the effective pull, $T_1 - T_2$. The decrease in shaft distance and the beginning of considerable slip, as $T_1 - T_2$ reaches values equal to the tensioning load, are clearly shown. This figure represents only one of a number of tests both with fan load and Prony-brake load, and with different speeds Only this one plot is given here because it represents runs at low and constant speed which made for accuracy in both Prony-brake and speed readings. At higher speeds the decrease in shaft distance came, however, even before the slack-side tensions approached zero. The graph should be taken mainly merely as a qualitative illustration on what takes place.

A horizontal drive with a nominal contact angle of 180 deg, represents the most favorable condition for transmitting large effective pulls $T_1 - T_2$, even with very low slack-side tensions. With smaller contact angles, such as occur in a three-point drive, the higher transmissive power obtainable with fixed shaft distance might not be so noticeable. Therefore this advantage had to be established by direct tests on such drives. The results will be reported in detail in the next section. They, however, showed that with the same initially applied tension, more effective pull could uniformly be transmitted with fixed shaft distance than with a floating-tensioning device. The advantage of the floating drive with tensioning device lies in compensating for permanent belt stretch.

Contact-Angle Apparatus

The determination of the influence of contact angle had proved impossible with fan load on the belt. With such a load, changes in the effective pull transmitted demanded changes in speed, and these changes introduced irregularities of air flow which scattered the data in such a manner that reputable curves expressing the relation between contact angle and effective pull transmissible could not be drawn. To get smooth and reputable curves it was found necessary to connect the driven shaft to a Prony brake by means of a flexible shaft, and in addition, run the tests at comparatively low speeds (about 500 to 600 rpm). It did not prove possible to find in the market a flexible shaft which could transmit the maximum torques desired for considerable intervals at high speed without excessive friction loss in the casing—hence heating and ultimate breakage.

The apparatus used is shown in Figs. 12 and 13. The driven shaft is carried in ball bearings at the lower end of a swinging arm. The pivot point of this arm at the upper end is carried in a sliding block, the location of which may be adjusted by means of a screw and handwheel, so that when readings were taken, the arm would hang vertically. Hence the total tension in the belt would be applied entirely by means of standard dead weights hanging as shown to the right in Fig. 12 and corrected for weight of weight carrier, etc. This view also shows the flexible shaft, the Prony brake, and the scale, by means of which the torque transmitted to the drive shaft could be read. The drive comes in from a dynamometer at the left in Fig. 12 to the lower test pulley. The arrangement of the test belt and the test pulleys is shown more clearly in Fig. 13. The upper pulley at the left is an idler. The contact angles could be changed by changing the positions of the the three shafts in the slots as indicated.

Four men were required to operate the device: Two to take simultaneous speed readings by means of Hassler tachometers at the driving and driven end, one to set the scale and keep the scale beam floating correctly during every test, and one to take down the readings and keep a running check on the percentage slip developed.

Loads were set on the scale from an arbitrary minimum up to a load at which the drive broke down, or at least developed a slip of 10 per cent or higher.

Plotting. The dead loads were transformed to $(T_1 + T_2)$ values by division by the cosine of the angle by which the test belt deviated from the horizontal. The $(T_1 - T_2)$ values, derived from the torque readings, were plotted against slip for various constant values of $(T_1 + T_2)$. Each plot applied to one contact angle at the driven end. When, for large contact angles at the driven end, the contact angle would have been smaller at the driving end, one large driving pulley was as a rule substituted at the driving end for the driven pulley and idler.

From the first plots so made, replots were made of $(T_1 - T_2)$ against contact angle for constant $(T_1 + T_2)$ values, and for a minimum and a maximum slip. Usually the minimum slip was 2 per cent; but for the largest pulleys (9 in. OD), it was 1 per cent; and for the smallest pulleys, in some cases, 4 per cent.

The pulley diameters were such as might be convenient for front-end automotive drives, but were smaller than those recommended for long life for industrial drives. Even in industrial drives, particularly with large transmission ratios, such small pulleys might also be found. Some manufacturers seem to get

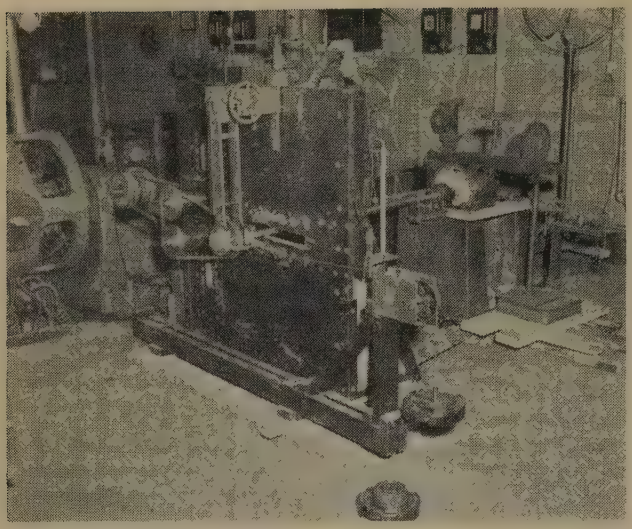


FIG. 12 CONTACT-ANGLE APPARATUS

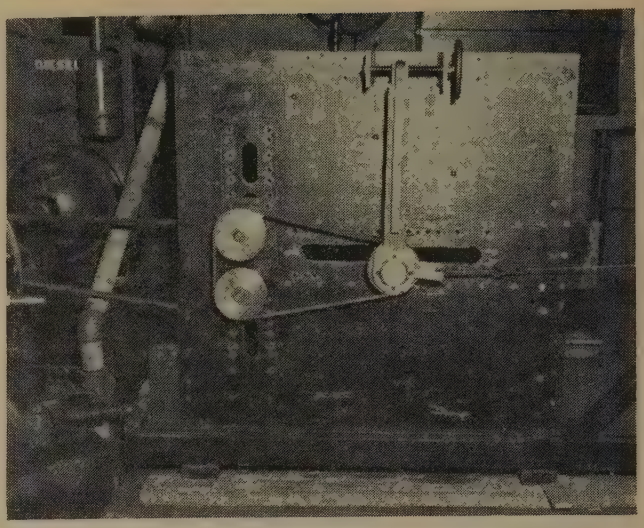
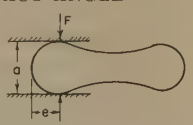


FIG. 13 FRONT VIEW OF CONTACT-ANGLE APPARATUS

TABLE 1 DEFICIENCY IN CONTACT ANGLE

$$K = \frac{Fea}{2} \qquad \cos C = 1 - \frac{K}{TR^2}$$



TYPE OF BELT	PITCH RAD. CUR. IN. (a/2)	F (LB)	K	C (DEG) (T=10 LB)	(T=5 LB)	(T=3 LB)	(T=1 LB)
C-68 COTTON COG ① e=3.50	2.50	1.05	9.2	31° 33'	45° 5'	59° 24'	118° 2'
	2.00	2.45	17.2	55° 15'	81° 57'	115° 41'	——
C-68 SMOOTH STEEL CABLE ② e=3.50	2.75	2.80	27.0	51° 19'	75° 31'	104° 29'	——
	2.25	5.40	42.5	80° 48'	132° 51'	——	——
C-68 STEEL CABLE COG ③ e=3.75	2.75	1.10	11.3	31° 44'	45° 30'	59° 52'	119° 40'
	2.00	2.50	18.8	58° 0'	86° 34'	124° 32'	——
B-60 STEEL CABLE COG ④ e=3.25	2.25	1.00	7.3	31° 8'	44° 36'	58° 40'	116° 6'
	1.50	2.25	11.0	59° 16'	88° 44'	129° 3'	——
B-60 SMOOTH COTTON ⑤ e=3.50	2.50	0.80	7.0	27° 23'	39° 6'	51° 10'	96° 54'
	1.75	2.55	15.6	60° 40'	91° 9'	134° 26'	——
FA-59 SMOOTH COTTON ⑥ e=3.25	2.00	0.50	3.25	23° 15'	33° 7'	43° 11'	79° 12'
	1.25	1.50	6.10	52° 29'	77° 24'	107° 38'	——
A-60 STEEL CABLE ⑦ e=3.25	2.00	0.40	2.6	20° 46'	29° 32'	38° 28'	69° 31'
	1.25	1.30	5.3	48° 37'	70° 37'	97° 28'	——
62" WEDGE RAYON COG 3/8" WIDE ⑧ e=3.25	2.25	0.20	1.5	13° 59'	19° 49'	25° 40'	45° 15'
	1.50	1.00	4.9	38° 33'	55° 40'	74° 9'	——
WEDGE STEEL CABLE COG 3/8" WIDE ⑨ e=3.25	2.25	0.20	1.46	13° 47'	64° 55'	87° 42'	——
	1.50	0.60	2.92	29° 32'	42° 16'	55° 28'	107° 27'

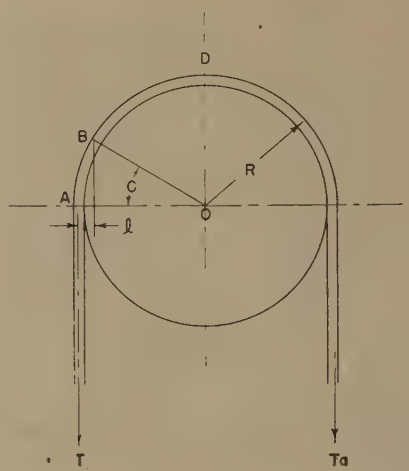


FIG. 14 CONTACT-ANGLE DIAGRAM

away with these by "over-belting" the drives, that is, introducing more belts and lower loads per belt than would be rational for larger pulleys.

In all these cases the question arises concerning the size of the actual contact angle. This angle is determined conventionally by representing the belt by straight lines drawn tangent to the pulley circumference. With any belt stiffness whatever this is always wrong, since at the running-on or running-off point of the belt the tension has no moment arm wherewith to flex the belt to the radius of curvature of the pulley. With pulleys as small as here used the actual contact angle might be very much smaller than the one determined by the tangents.

To get a rough idea of what the deviation might amount to, test belts were bent double between the platens of a testing machine, as shown by the figure at the top of Table 1. For a wire bent elastically the bending moment is equal to K/R, where K is a proportionality factor, and R is the radius of curvature to which the wire is bent. With the notation used at the top of Table 1, we would have $K = Fea/2$, F being the bending force exerted by the testing machine.

On the other hand, if with reference to Fig. 14, C is the angle at which the belt conforms to the pulley radius R, we have $Tl = TR\,(1 - \cos C) = K/R$. Hence as indicated at the top of Table 1

$$\cos C = 1 - \frac{K}{TR^2} \qquad [6]$$

By means of this formula the C values listed in Table 1 were obtained. They apply to pulley pitch diameters of from $2^1/_2$ to $5^1/_2$ in. (pitch radii of 1.25 to 2.75 in). It will be seen that on the slack side, where tensions may well go down to the low values of 1 to 10 lb listed, the deficiency angles C may be so large that it

becomes meaningless to use theoretically derived formulas like Equation [1] where the contact angle is taken to be the one determined by tangents in the conventional fashion. On the other hand, correction formulas like the one at the top of Table 1 are also not of much practical value for computation, since the stiffness constants, though they may be used as here for a very rough orientation, are not usually known for belts to be used and cannot be determined simply with any accuracy.

The only practical solution seems to be to express, if possible, in some simple way the actual $(T_1 - T_2)$ values obtained for various nominal contact angles with various belts and various $(T_1 + T_2)$ values.

Results. The first test results to be plotted were for C-size belts on 6.9-in., 5-in., and 4-in-OD pulleys. It seemed as if these results could be represented fairly by a linear relation for which was suggested the form

$$T_1 - T_2 = \frac{k\alpha}{100}(T_1 + T_2) \quad \ldots\ldots [7]$$

α being the conventional contact angle in degrees, and k a constant to be determined from the plot. All these pulley diameters are below the absolute minimum of about 9 in. OD recommended for long-life industrial drives. Tests with C-size belts on 9-in-OD pulleys showed hardly any influence of contact angle at all between angles of 90 deg and 180 deg. It should be noted, however, that these tests were run with effective and total pulls considerably below the rated values for these belts, since the full rated torques on 9-in-OD pulleys would have exceeded the capacity of the flexible shaft.

If any conclusion is to be drawn from this it would be that the results were due to the comparatively long arc length of contact on large pulleys, as well as to the large radius of curvature and consequently lessened influence of belt stiffness in reducing contact angle.

Strange to say, a similar independence of contact angle may be observed also on very small pulleys, Fig. 20. The reason then may be that belt stiffness reduces the actual contact almost to a point for all angles. However, against this, see under "Addendum", page 343.

If we assume, as we must, that $T_1 - T_2$ is zero for $\alpha = 0$, it would seem that the best compromise formula to suggest is

$$T_1 - T_2 = k\sqrt{\frac{\alpha}{100}}(T_1 + T_2) \quad \ldots\ldots [8]$$

As Figs. 15 to 20, inclusive, show, this formula represents acceptably the relation for B-size belts on 5 in. OD with practically interesting loads; and does it for both smooth cotton V-belts and, steel-cable notch belts, and for fixed-shaft-distance ("locked" drives), as well as for floating-shaft-distance ("unlocked" drives). For a comparatively stiff smooth cotton belt on pulleys of only $3\frac{1}{2}$ in. OD, there is a tendency for the experimental plots to run flat.

Comparison With Experimental T_1/T_2 Values. From the practical point of view a formula giving the effective pull $T_1 - T_2$ directly from the total pull $T_1 + T_2$, is decidedly preferable to a formula giving T_1/T_2. The use of T_1/T_2 as a belt-drive characteristic springs merely from the theoretical formula

$$T_1/T_2 = e^{f\alpha/\sin\gamma}$$

Since values of the coefficient of friction f must be derived from tests which give $T_1 - T_2$, and $T_1 + T_2$ directly, and must be derived under assumption of a contact angle, which may have little to do with the actual contact angle, its use is very much like building a mountain merely to be able to tunnel through it. However, as the author, in the past, has been an advocate of T_1/T_2 values, determined experimentally, as a belt characteristic, and this usage has found widespread acceptance, it may be well to compare T_1/T_2 values derived from Equation [8] with values now commonly used in industry.

It can readily be seen that

$$\frac{T_1}{T_2} = \frac{(T_1 + T_2) + (T_1 - T_2)}{(T_1 + T_2) - (T_1 - T_2)} \quad \ldots\ldots [9]$$

The highest value for T_1/T_2 which could be conceived would be infinity; this would be obtained for $T_2 = 0$. Actually, the author has seen V-belt drives with 180 deg nominal contact angle in which T_2 very closely approximated this value. We would then have $T_1 - T_2 = T_1 + T_2 = T_1$.

Equation [8] would take the form

$$T_1 = k\sqrt{\frac{\alpha}{100}}\,T_1$$

from which

$$k = \sqrt{\frac{100}{\alpha}} = 0.745$$

At a contact angle of 100 deg, $\alpha/100$ is equal to unity, and Equation [9] would with $k = 0.745$ give

$$\frac{T_1}{T_2} = \frac{1.745\,(T_1 + T_2)}{0.255\,(T_1 + T_2)} = 6.85$$

A k-value of 0.745 is evidently the maximum possible for floating drives. If higher values are obtained from smoothed curves there may be two reasons, i.e., either that experimental readings were not accurate enough, or that the smoothed curves were drawn under too much influence of points for low contact angles.

In Table 2 are listed values for k. The symbol k_1 stands for floating drive; k_2 stands for locked drive. If the total tension $T_1 + T_2$ refers to a tension set by dead weights before the center distance was locked, k_2 can well be above 0.745 for reasons set forth in the section on driving action. It will be observed, however, that excessively high k_1 values are also listed. In case of the small A-size belts, this, no doubt, is due mainly to inaccuracies of scale readings with the mall torques transmitted. In the case of the C-size belts, the reason is rather to be seen in smoothed cu ves drawn too high

It is greatly to be regretted that there is now no possibility of correcting these defects. They were accumulated during the author's illness. It is possible that the data in Table 2 should not be presented at all. Nevertheless, if they are presented, it is for one reason only, namely, that k_2, the constant for locked-shaft distance, is uniformly higher than k_1, the constant for floating drive. That this was so the author can testify from all the tests he witnessed, personally, at which he often did the locking himself.

The values of k_1 and k_2 to use in computation practice, for the present must be settled more by recourse to general practical experience than to the data in Table 2.

Let us begin by taking the k value of 0.632, which was found for B-60 belts both of steel-cable notch and of smooth cotton construction.

This corresponds at 180 deg contact angle to a T_1/T_2 value of 12.3; and at 100 deg to one of 4.45. These are values which well may be attained, although for new drives somewhat lower values usually are assumed.

A value of $k_1 = 0.6$ would for 180 deg give $T_1/T_2 = 9$, and for 100 deg give one of 4. These are good, but quite attainable values.

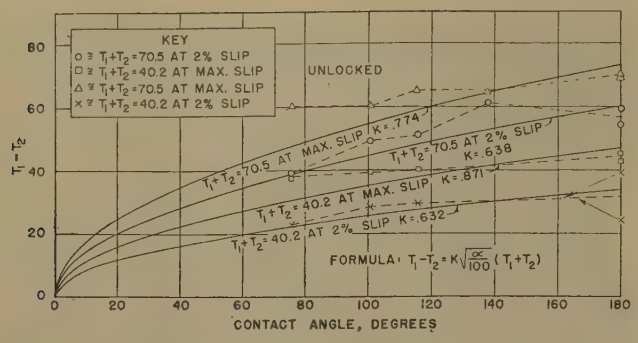


FIG. 15 EXPERIMENTAL CONTACT-ANGLE CURVES, USING GOODYEAR B-60 SMOOTH COTTON V-BELT, 5 IN. NOMINAL PULLEY OD, AND POINTS FOR SLIP TAKEN AT 2 PER CENT AND AT MAXIMUM; DRIVE UNLOCKED

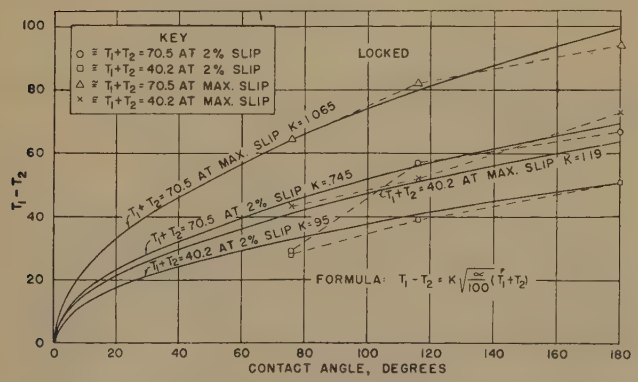


FIG. 16 EXPERIMENTAL CONTACT-ANGLE CURVES, USING GOODYEAR B-60 SMOOTH COTTON V-BELT, 5 IN. NOMINAL PULLEY OD, AND POINTS FOR SLIP TAKEN AT 2 PER CENT AND AT MAXIMUM; DRIVE LOCKED

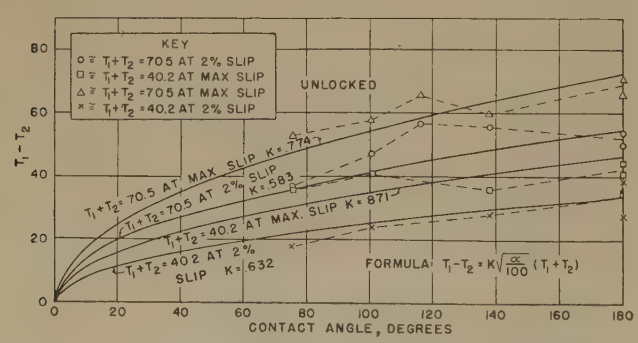


FIG. 17 EXPERIMENTAL CONTACT-ANGLE CURVES WITH GOODYEAR B-60 STEEL-CABLE NOTCH V-BELT, 5 IN. NOMINAL PULLEY OD, AND POINTS FOR SLIP TAKEN AT 2 PER CENT AND AT MAXIMUM; DRIVE FLOATING

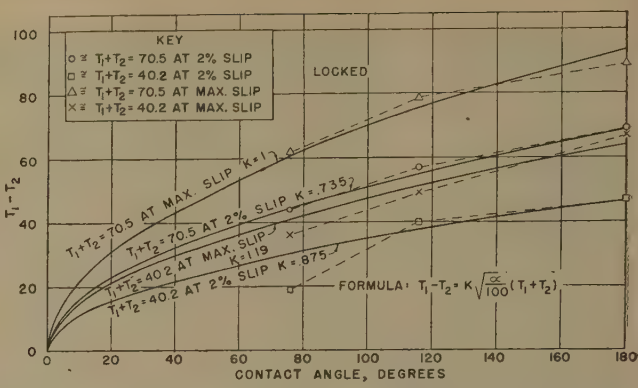


FIG. 18 EXPERIMENTAL CONTACT-ANGLE CURVES WITH GOODYEAR B-60 STEEL-CABLE NOTCH V-BELT, 5 IN. NOMINAL PULLEY OD, AND POINTS FOR SLIP TAKEN AT 2 PER CENT AND AT MAXIMUM; DRIVE LOCKED

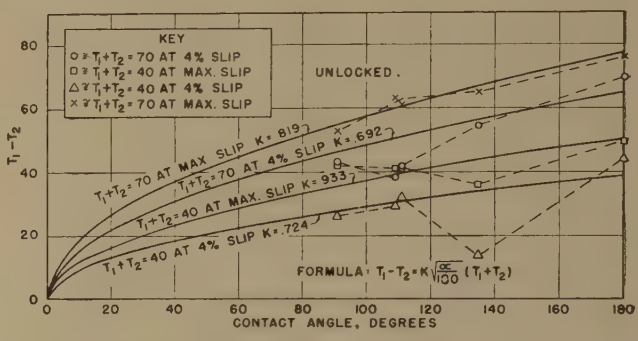


FIG. 19 EXPERIMENTAL CONTACT-ANGLE CURVES, WITH GOODYEAR B-60 STEEL-CABLE NOTCH V-BELT, 3.5 IN. NOMINAL PULLEY OD, AND POINTS FOR SLIP TAKEN AT 4 PER CENT AND AT MAXIMUM; DRIVE FLOATING

(Points at contact-angle 109° 8′ for driver pulley and all others for driven pulley.)

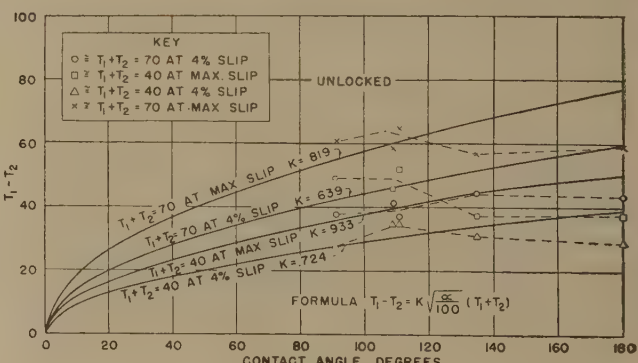


FIG. 20 EXPERIMENTAL CONTACT-ANGLE CURVES, WITH GOODYEAR B-60 SMOOTH COTTON V-BELT, 3.5 IN. NOMINAL PULLEY OD, AND POINTS FOR SLIP TAKEN AT 4 PER CENT AND AT MAXIMUM; DRIVE FLOATING

(Points at contact-angle 100° 8′ for driver pulley and all others for driven pulley.)

TABLE 2 VALUES OF K

$T_1 - T_2 = K\sqrt{\frac{\alpha}{100}}\,(T_1 + T_2)$

K_1 = CONSTANT FOR UNLOCKED CENTER DISTANCE
K_2 = CONSTANT FOR LOCKED CENTER DISTANCE

TYPE OF BELT	PULLEY DIAM IN.	K_1	K_2	K_2/K_1	K_1	K_2	K_2/K_1	K_1	K_2	K_2/K_1	REMARKS	
		$T_1+T_2=20$			$T_1+T_2=30$			$T_1+T_2=50$			% SLIP	t/OD
FA-59 SMOOTH COTTON	4				0.795	0.914	1.150	0.718	0.922	1.383	2	0.0781
FA-59 SMOOTH COTTON	6	0.968	1.115	1.152				0.796	0.886	1.112	2	0.0521
A-60 STEEL CABLE COG	4				0.770	0.891	1.156	0.717	0.792	1.105	2	0.0860
A-60 STEEL CABLE COG	6	0.844	0.959	1.135				0.732	0.834	1.139	2	0.0573
		$T_1+T_2=40$			$T_1+T_2=70$.							
B-60 SMOOTH COTTON	3-1/2	0.724			0.639						4	0.116
B-60 SMOOTH COTTON	5	0.632	0.950	1.500	0.638	0.745	1.170				2	0.0813
B-60 SMOOTH COTTON	6-1/4	0.527			0.479						2	0.065
B-60 STEEL CABLE COG	3-1/2	0.724			0.692						4	0.125
B-60 STEEL CABLE COG	5	0.632	0.875	1.380	0.583	0.735	1.260				2	0.0875
B-60 STEEL CABLE COG	6-1/4	0.615			0.585						2	0.070
		$T_1+T_2=46$			$T_1+T_2=121$							
C-68 SMOOTH STEEL	4	0.400			0.400						2	0.1408
C-68 SMOOTH STEEL	5	0.500			0.500						2	0.1125
C-68 STEEL CABLE COG	4	0.400			0.400						2	0.1408
C-68 STEEL CABLE COG	5	0.500			0.500						2	0.1125
		$T_1+T_2=39$			$T_1+T_2=90$							
C-68 SMOOTH STEEL	7	0.600			0.600						2	0.0805
C-68 STEEL CABLE COG	7	0.600			0.600						2	0.0805
		$T_1+T_2=30$			$T_1+T_2=40$			$T_1+T_2=50$				
C-68 STEEL CABLE COG	9	0.744						0.744			1	0.0625
C-68 NO. 15 SMOOTH COTTON	9	0.795	1.015	1.275				0.671			1	0.0625
C-68 NO. 16 REG. COTTON COG	9	0.810	0.945	1.167	0.721	0.821	1.139	0.696	0.746	1.072	1	0.0625
WEDGE BELTS		$T_1+T_2=40$			$T_1+T_2=50$							
STEEL CABLE COG - 3/8" WIDE	4	0.616	0.725	1.178	0.583	0.715	1.228				2	0.1093
RAYON - 3/8" WIDE	4	0.650	0.745	1.148	0.660	0.765	1.159				2	0.1093

A conservative value would be $k_1 = 0.5$, corresponding for 180 deg to $T_1/T_2 = 5$, and for 100 deg to $T_1/T_2 = 3$.

The lowest k_1 value listed is 0.4, which would give $T_1/T_2 = 3.34$ for 180 deg and 2.34 for 100 deg. This is distinctly too low for V-belts, but would be found not rarely for flat belts.

A designer who wishes to play perfectly safe even for small pulleys may take $k_1 = 0.5$. With reasonably large pulleys, k_1 values of 0.6 might be attained readily.

Locked Versus Floating Drives. In Table 2 are listed values for k_2/k_1, k_1 being the k-value for floating drive, and k_2 the value for fixed shaft distance. It will be seen that in only one case is k_2/k_1 less than 1.1, while in many cases it is considerably higher. Hence it is conservative to say that at least a 10 per cent higher effective pull can be transmitted with fixed shaft distance than with a floating drive, the initially set $T_1 + T_2$ value being the same; provided of course permanent stretch is still inconsiderable.

It should be noted that where the shaft distance and the belt tightness are readily adjustable, it is advisable to lock the shafts while the drive is idling slowly, since then there is assurance of an even tension all around the drive.

In the experiments here reported the locking was accomplished by applying to the tensioning cable a clamp butting against the dead-load pulley stand at the right in Fig. 12. As a rule this was done with the drive idling slowly.

However, locking at standstill was tried and, in the author's opinion, seemed to give less consistent results.

Addendum. The thought may occur that on very small pulleys the arc of contact with belts of some stiffness would be so small that it could be regarded as merely point contact. In such a case it can be seen readily that the total force pressing the belt against the pulley would be $(T_1 + T_2) \sin \alpha/2$, where α is the conventional angle of contact. The friction force preventing slippage would be

$$\frac{f}{\sin \gamma}\,(T_1 + T_2) \sin \frac{\alpha}{2},$$ where γ is one half the groove angle,

and we might set

$$T_1 - T_2 = \frac{f}{\sin \gamma}\,(T_1 + T_2) \sin \frac{\alpha}{2}$$

With $f = 0.25$ and $\gamma = 14^1/_2$ deg, $f/\sin\gamma$ would be unity. For $\alpha = 180$ deg, we would have $T_1 - T_2 = T_1 + T_2$, i.e., the slack-side tension T_2 would be zero, and T_1/T_2 would be infinity.

For a contact angle of 90 deg, $\sin \alpha/2$ would be 0.707, and T_1/T_2 would be 8.15.

On the other hand, if $f/\sin\gamma$ were 0.8, T_1/T_2 for $\alpha = 180$ deg would be 9, and for 90 deg it would be 3.6. These might be quite reasonable values for practical use.

Pulley Crown

Apparatus. The pulley-crown tests at The Ohio State University had been preceded by so-called crown-conformation tests at Akron. These indicated that wide modern belts, particularly with steel-cable transmitting organs, might not conform to the pulley rim at the edges, unless the crown was quite low. This of course would have a tendency to concentrate the power transmission of the belt toward its middle section only, which could not have a favorable effect on the life of the belt.

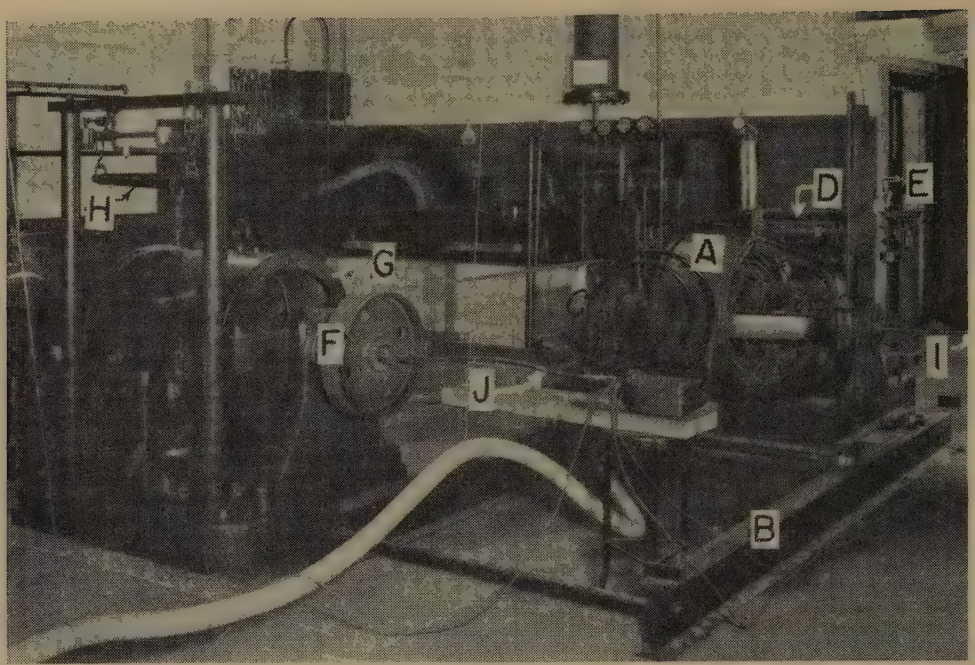


FIG. 21 PULLEY-CROWN-TEST APPARATUS
(*A*, Cradled motor; *B*, grooved I-beams; *C*, pulley; *D* and *E*, weighing scale; *F*, pulley; *G*, dynamometer; *H*, scale arm; *I* and *J*, speed-indicator cables.)

The question then arose how much crown was absolutely necessary to prevent the belt from running off the pulleys.

Tests on this were performed in the laboratory of the mechanical-engineering department of The Ohio State University with the apparatus shown in Fig. 21. In this, *A* is a motor cradled on a stand so that it can oscillate, making it possible to weigh the torque on scale D. The belt runs from a pulley *C* on the motor to a pulley *F* on a dynamometer *G*, on the scale arm *H* of which the torque at the driven end can be weighed. The cradled motor ran on grooved I-beams *B*, and the total pull on the belt could be weighed through a cable and a bell crank, bearing on a platform scale attached to the far end of the left I-beam.

The whole I-beam frame could be shifted on the floor so as to produce various angles of misalignment between the motor shaft and the dynamometer shaft. The angularities were determined by markings on the floor. By good luck, the help of Mr. O. J. Marshall, then professor of geodetics and surveying at The Ohio State University, now at the University of Toronto, was secured for these angle determinations. His painstaking care and systematic orderliness in this matter cannot be too highly appreciated.

Belts Used. The test belts used were the following:

No. 1 3 in. × 16 ft endless steel-wire Compass No. 250.
No. 2 3 in. × 9 ft endless steel-wire Compass.
No. 3 3 in. × 16 ft endless cotton-cord Compass No. 40.
No. 4 3 in. × 16 ft open-end, 5-ply fabric Thor (laced with Clipper fasteners).
No. 5 10 in. × 16 ft steel-wire Compass No. 250.
No. 6 10 in. × 16 ft endless cotton-cord Compass No. 40.
No. 7 10 in. × 16 ft 5-ply fabric Thor open-end (laced with Clipper fasteners).

The pulleys used were, in the majority of cases, cast-iron ones, 11-in. wide, and of $8^3/_4$ in. and 22 in. diam over the crowns. This gave a transmission ratio of about $2^1/_2$. Various heights of crown were secured by successive turning down of the pulleys to increasing taper from the central ridge, the diameter of which remained constant

All crowns are expressed as excess of the diameter at the central ridge over the diameter at the edges.

Results. Initial orientation tests showed that a crown of 0.06 in. on the diameter per foot of pulley width was sufficient to steer the belt, but only if the shafts were specially aligned for each belt. A crown of 0.10 in. per ft was necessary to guide all belts without special alignment.

In Fig. 22 is given a general plot of the results of all tests. The tests here plotted were all idling tests with varying degrees of total tension $T_1 + T_2$. However, runs with actual power transmission and unvarying $T_1 + T_2$ were carried out to make sure that the belts still tracked properly. The capacity of the motor-dynamometer aggregate was, however, not sufficient to permit full rated load for the 10-in-wide belts.

The plots in Fig. 22 all refer to a speed-down drive, the motor speed being approximately 680 rpm, and the dynamometer speed 280 rpm. This meant a belt speed of about 1600 fpm. Speed-up drives with the 22-in. pulleys at the motor end were, however, also tried. This gave a belt speed of about 4000 fpm.

When the high-speed results were plotted over the low-speed results, the overlapping was such as to cause confusion, and one was tempted to conclude that speeds up to 4000 fpm had no appreciable effect on the guiding power of the crown.

However, separate plots of the speed-up tests, as given in Fig. 23, indicate that with high speed there is less influence of total tension on the results than at low speed, at least for the wider belts.

The angles of misalignment plotted in Figs. 22 and 23 were the angular change while the belt moved from one edge of the pulley to the other. A reasonable assumption would be that one half of this angle was deviation of the shafts from strict parallelism.

Narrow Belts. Since all tests were run on pulleys 11 in. wide, the observed possible deviations were greater for the belts of 3-in. width than those which in practice would be possible with pulleys not more than 1 in. wider than such belts. For this reason tests were also run with the 3-in. belts in which the observed angular change was the one occurring while the belts moved from a posi-

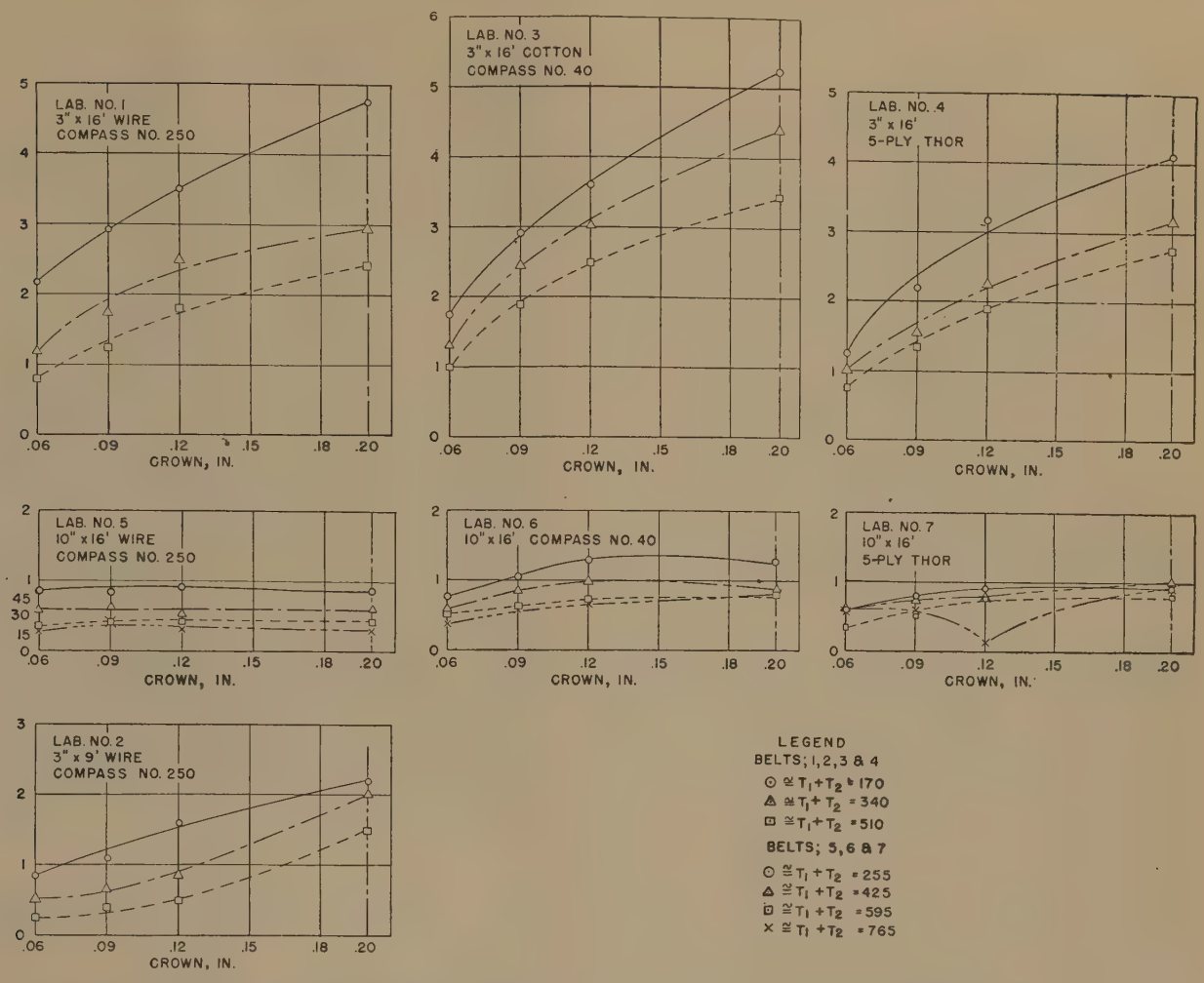


FIG. 22 PULLEY-CROWN TESTS WITH SPEED-DOWN DRIVE, USING 8¾-IN-DIAM MOTOR PULLEY
(Average motor speed 680 rpm; 22-in-diam dynamometer pulley; average dynamometer speed 280 rpm.)

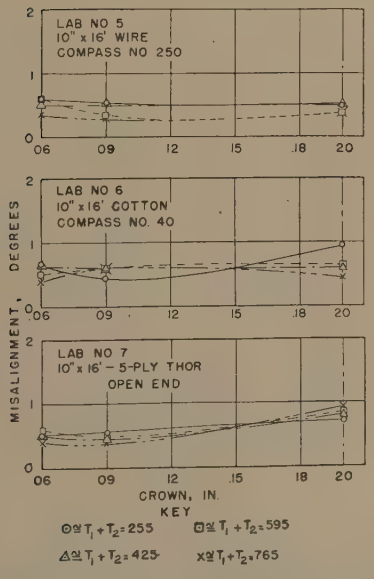


FIG. 23 PULLEY-CROWN TESTS WITH SPEED-UP DRIVE, USING 8¾-IN-DIAM DYNAMOMETER PULLEY, AND 22-IN-DIAM MOTOR PULLEY

tion ½ in. to the right of the central position to one ½ in. to the left. The results are plotted in Fig. 24.

It also will be seen in these tests that narrow belts show more sensitiveness to increased height of crown than do wide belts; but the maximum angles of permissible misalignment are not much greater than for wide belts under the same speed conditions (680 rpm at driving end, 280 rpm at the driven end).

General Observations:

1 Response to crown improves as tension decreases (at least in the speed-down drives, although this is not as pronounced for wide belts as for narrow belts).

2 Reponse to crown improves as ratio of length over width increases.

3 Steel belts and cotton belts respond roughly to the same extent to given crowns. (Note cases 1, 3, and 4, and cases 5, 6, and 7. Reservations might be made for cases 1 and 5.)

4 As already stated, increased height of crown has less effect on wide belts than on narrow belts. In view of the fact that wide belts may not conform to the pulley at the edges, it is well to use as small crowns as possible with wide belts.

5 Great care should be taken to align shafts properly for wide belts, since increased crown is not a remedy for poor alignment, and the belt may suffer from too much crown.

6 Since low tension increases the effectiveness of the crown,

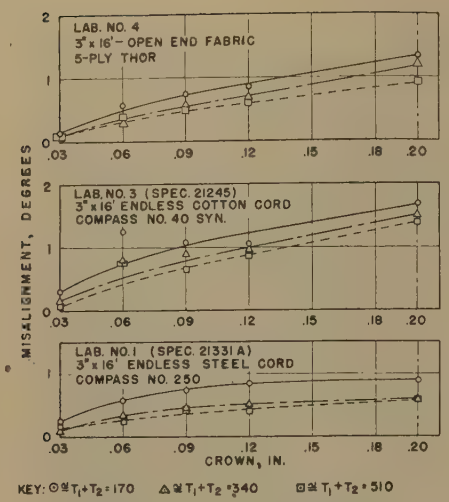


FIG. 24 PULLEY-CROWN TESTS WITH SPEED-DOWN DRIVE, USING 3-IN. BELTS ON SIMULATED 4-IN. PULLEY WIDTH

(Belt travels 1/2 in. from center, 8 3/4-in-diam pulley on motor; average motor speed 680 rpm, and 22-in-diam pulley on dynamometer; average dynamometer speed 280 rpm.)

it might be concluded that the crown is more effective on the driven pulley, on which the belt runs on slack, than on the driving pulley where it runs on tight. Some European belt practice seems to be based on this idea. Direct tests at The Ohio State University on crowning at one end only, however, rendered this conclusion questionable.

Theoretical Explanations. The observed behavior of the belts can be explained by the commonly given theory of crown effect. It is assumed that when the belt runs onto a tapered pulley, it flexes sideways as shown in Fig. 25, i.e., contact is made at a point nearer the high end as the belt runs on than is the simultaneous point at the center of belt contact. The wider the belt, the more difficult it is for it to flex sideways. The higher the pull, the greater is the tendency to pull the belt straight and reduce the possibility of sidewise bending.

A wider belt would not then be guided by a crown as effectively as a narrow belt, and a heavily loaded belt would not be guided as effectively as a lightly loaded one. Both conclusions are in harmony with the test results here reported.

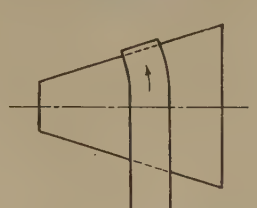


FIG. 25 CROWN ACTION

LIFE TESTS

Throughout the entire period during which the investigations reported in this paper were in progress, life tests on the belts of various manufacturers were run continuously on the author's differential life-testing machine for V-belts.[10] These tests gave important information as to what load should be regarded as normal for various types and sizes of belts.

Of course results cannot here be given, since they might in some cases be held detrimental to the interests of industrial concerns. However, one thing can be emphasized, and that is, the importance of most careful arrangement of the carrying elements of the belt section.

Of particular interest were also results with the so-called Wedge belts, in the development of which Mr. R. S. Carter played an important part, and which now have been applied as standard equipment on at least one well-known automobile. In these belts the V-section runs to a sharp edge as in the letter V itself, and is not truncated to give a flat inner surface.

SUMMARY OF FINDINGS

TESTS ON V-BELT DRIVES

Centrifugal Force. Centrifugal force may have much less influence than commonly assumed. This is true particularly for belts with small elastic stretch, such as belts with steel-wire transmitting organs; but is fairly true also for other modern V-belts, except that belts with textile transmitting organs may slip somewhat more than steel-wire belts.

Fixed Versus Floating Shafts. A drive with fixed shaft distance, as long as the belt has not undergone much permanent stretch, can transmit more torque than a floating drive with total tension set by a tensioning device.

Angle of Contact. On comparatively small pulleys, the actual arc of contact between a V-belt, or other stiff belts, and the pulley may be very much smaller than the conventional one obtained by representing the belt by straight lines drawn tangent to the pulley circumference. Therefore the actual effect of conventional contact angle can be obtained only by direct test. In the author's opinion, it may be represented most conveniently by a constant k in a formula

$$T_1 - T_2 = k\sqrt{\alpha/100}\,(T_1 + T_2)$$

where α is the conventionally determined contact angle in degrees, and T_1 and T_2 are the tight and slack-side tensions, respectively.

If k_1 is the constant for a floating drive, and k_2 the constant for the same drive with fixed shaft distance, the advantage of the latter can best be indicated by the ratio k_2/k_1. This ratio is almost always in excess of 1.1.

Recommended Pulley Diameters. From the point of view simply of the torque transmitted, the tests seem to indicate that for V-belts the ratio of pulley diameter to belt thickness should not be less than 12 (ratio of belt thickness to pulley diameter should not be more than 0.08). For so-called Wedge belts the diameter/belt-thickness ratio may be 10 (the inverse ratio 0.1). These values correspond also to the minimum diameters usually prescribed from the point of view of reasonable belt life.

Ratio of Total Tension to Initial Tension. The ratio of total tension with power transmission to total tension without power transmission initially applied depends, among other things, on the stress-strain curve of the belt. From this point of view the running total tension may be either smaller than, or larger than, or equal to the initial tension. If the slack-side tension drops to zero, so that the belt no longer can contract on the slack side, and if power is still transmitted, the total tension becomes simply equal to the tight-side tension.

CROWNING OF FLAT-BELT PULLEYS

Pulley Crown. Modern flat belts, particularly with steel-wire transmitting organs, do not conform to a crown as well as old-time leather belts. It is bad for the belt to be riding merely on a ridge and to transmit the full tension only in the middle. Hence the necessity for finding the minimum crowns needed to keep the belts guided.

[10] U. S. Patent No. 1749297; "An Endurance Testing Machine for Belts," Circular No. 16, Engineering Experiment Station of The Ohio State University; also advertising material of Tinius Olsen & Company, Philadelphia, Pa.

It was found that a crown of 0.10 in. on the diameter per foot of pulley width was necessary to guide all belts without special adjustment of shaft alignment.

A crown of 0.06 in. will be satisfactory for modern belts, if the shafts are specially aligned for the belts used. Narrow belts respond to increased crown more readily than wide belts. Slack belts are more easily guided than belts under heavy tension.

Acknowledgments

Indebtedness to Messrs. Carter and Suloff and to the Goodyear Tire and Rubber Company as such has already been acknowledged. To that company and to those gentlemen go the credit for the inception, the financial support, and the unvarying determination to see the investigation through to a reasonable conclusion.

Here the writer wishes to acknowledge his profound indebtedness to those who in one way or another participated in the test and recording work at The Ohio State University. Prof. O. J. Marshall has already been mentioned in connection with crown tests; but he took part also in a great deal of V-belt testing. Prof. George N. Moffat did very helpful work in elastic-stretch determinations, and in emergency test supervision during the author's illness.

Indispensable aid was continuously rendered by Mr. Roy L. Pratt, stationary engineer in the mechanical-engineering department of The Ohio State University, both in assembling apparatus and in helping to run it at tests. Similar aid was rendered by the late Mr. R. T. Simpson in the manufacture of apparatus. Of younger faculty members who at one time or another participated in the work, the names of Messrs. Lee Toliver, D. H. Whiston, and G. D. Hudelson should be remembered.

Any attempt to list names of the numerous students who helped in testing and computation would be unjust since inadvertently some names would be omitted. Nevertheless, the author would like to mention Messrs. T. E. Swander, R. B. Clapper, and H. A. and C. W. Porterfield. The typing help of Mrs. Ruth M. Manker and Miss Arlene Schantz was indispensable in getting out the manuscript.

Finally, the courtesy and co-operation of Prof. J. R. Shank, Mr. H. J. Hoffman, and Mrs. Irene E. Harris of the staff of the Engineering Experiment Station of The Ohio State University are gratefully remembered in connection with the final preparation of a Bulletin.

Supplementary Tests

At the time of the final proof reading of this paper the Goodyear Company has made it possible to run certain additional tests with an automotive propeller shaft substituted for the flexible shaft shown in Fig. 12.

With this equipment C-68 steel-wire and cotton belts could be run up to rated loads on 6.9 OD sheaves at belt speeds of 4000 to 5000 ft. per min.

The results so far obtained indicatethat the k values in formula [8] derived from runs at lower speeds and loads, apply also at these higher speeds and loads; and that, as before found, higher k values are permissible for fixed shaft distance than for floating drives.

This additional work was done by Mr. Robert E. Dine with the assistance of Mr. R. L. Pratt and students.

Anisotropic Plastic Flow

By J. C. FISHER,[1] SCHENECTADY, N. Y.

The mathematical theory of plasticity has been generalized to apply to plastically anisotropic materials. The anisotropic flow theory has been shown to reduce to the resolved shear stress - shear strain relationship for single crystals, and to the distortion energy theory for isotropic polycrystals. Experimental data concerning the plastic flow of anisotropic polycrystalline aluminum indicate that the predictions of the anisotropic flow theory are in good agreement with experiment, and are significantly better than the predictions of the distortion energy theory. As plastic anisotropy is the rule rather than the exception, caution is indicated in interpreting the results of combined stress tests by means of the distortion energy theory.

THE fundamental problem in the field of plasticity is that of determining the plastic strains and strain rates corresponding to an arbitrary history of the temperature and stresses applied to an element of material. The mathematical theory of plasticity generally restricts itself to deformations at constant temperature and constant rate, and discusses the relationships between strain and stress with these limitations. Subject to the constant-temperature and constant-rate conditions of the mathematical theory of plasticity, stresses and plastic strains have been related successfully for the deformation of both single crystals and isotropic polycrystals. However, the theories for these two cases appear to have little in common, and they have not led to a theory for the general case of plastic flow in an anisotropic polycrystal.

It is desirable to develop a generalized theory of crystal plasticity which will reduce under the appropriate conditions to the accepted single-crystal and isotropic-polycrystal treatments, and which in addition will describe the plastic flow of anisotropic polycrystals. The present paper attempts such a generalization of the theory of plasticity, on the basis of the slip mechanism of plastic deformation common to both single-crystal and polycrystalline materials.

Basic Theory—Nucleation of Slip

Observation of plastic flow in single crystals has led to the discovery of a number of phenomena associated with the flow process (1, 2).[2] A few of the more important of these, from the standpoint of plasticity theory, are summarized as follows:

1 Plastic flow frequently occurs by slip, a relative motion of two portions of a crystal where the two portions slide past one another for many (the order of hundreds) atomic spacings.

2 The relative displacement observed during slip occurs on certain crystallographic planes and not on others.

3 The relative displacement is in a particular crystallographic direction.

4 At a constant temperature and shear strain rate there is a unique relationship between the component of shear stress in the slip plane and in the direction of slip (called the resolved shear stress) and the plastic shear strain corresponding to this stress (called the resolved shear strain), which holds for plastic deformation of a given crystal under arbitrary stress systems.

Tests of polycrystalline metals also indicate that plastic deformation proceeds by slip, with the extent of each slip band limited to a single crystal of the matrix. Slip frequently occurs simultaneously on different sets of planes in adjacent regions of the same grain, giving rise to deformation bands (2). Although the resolved shear stress - shear strain relationship loses its meaning for polycrystals, the von Mises-Hencky distortion energy theory of plastic flow leads to a unique effective stress-effective strain relationship for isotropic polycrystals (3, 4).

Slip bands may be thought of as regions of relative motion between two plastically undisturbed portions of a crystal. They grow with extreme rapidity, the time for complete formation being of the same order as for the passage of sound through a similar distance, as indicated by the clicking noises accompanying slip in rock salt (5). During the slip process, slip bands appear unable to support a shearing stress in the direction of slip. After slip has occurred, however, the bands very rapidly regain their ability to withstand shearing stresses.

It is reasonable to believe that during slip the elastic strain energy near the slip band is relieved to an extent consistent with the limited relative displacements during the process. For a slip band with area A, elastic energy should be relieved in a volume roughly equal to that of a sphere with great circle area A, as indicated schematically in Fig. 1. A smaller amount of strain energy should be concentrated near the perimeter of the slip band (unless the band terminates at a free surface of the metal) as sketched in the same figure.

Measurements of the elastic energy stored in a deformed poly-

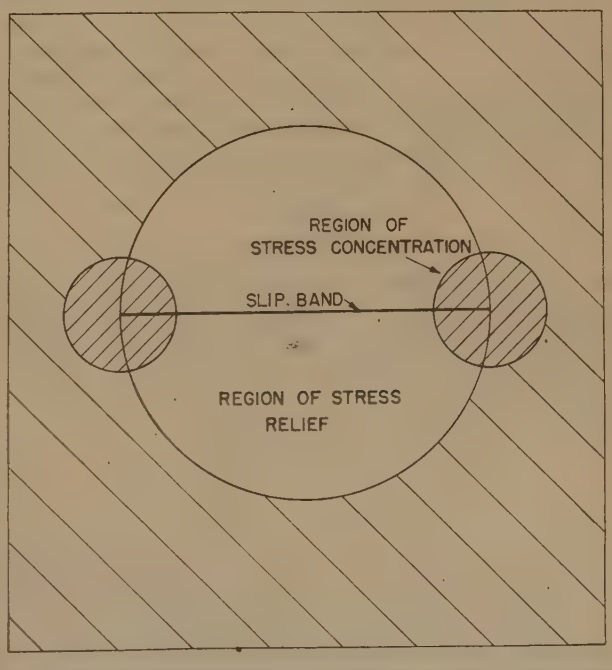


Fig. 1 Stress Distribution in Neighborhood of a Slip Band, Schematic

[1] Research Associate, General Electric Research Laboratory. Mem. ASME.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Metals Engineering Division and presented at the Semi-Annual Meeting, Milwaukee, Wis., May 30–June 5, 1948, of The American Society of Mechanical Engineers.

crystal indicate that on the average the strain energy concentrated near the perimeter of a slip band is only about 10 per cent of the energy relieved by the slip process (6). Therefore, the stress system in the neighborhood of a newly formed slip band will resemble that shown in Fig. 2, where the energy concentrated

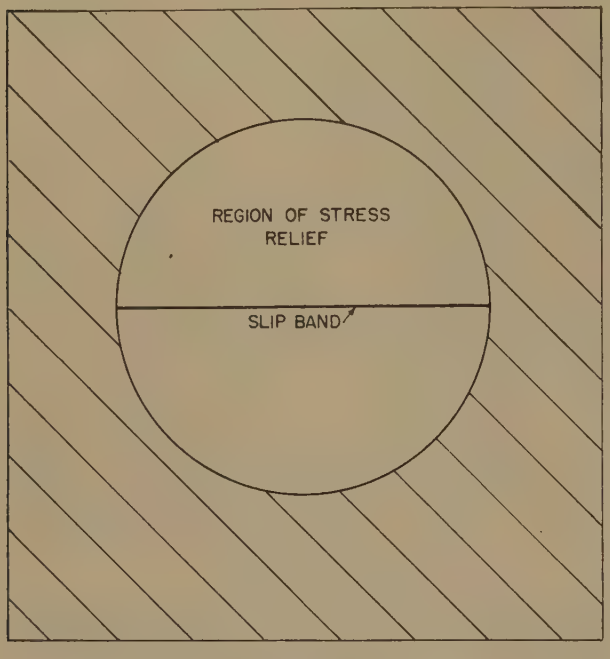


FIG. 2 STRESS DISTRIBUTION IN NEIGHBORHOOD OF A SLIP BAND, SCHEMATIC

at the perimeter of the band has been neglected and that portion of the strain energy relievable by the slip process has been relieved in a volume containing the band.

Statistical mechanical considerations require that very small slip bands appear and disappear from time to time in a stressed crystal. Most of these bands are smaller than a critical minimum size required for continuous growth, and will be called "slip embryos." A few occasionally exceed the critical size, become "slip nuclei," and grow rapidly as observable slip bands.

The problem of slip band nucleation has been discussed by Leschen and Carreker (7), who assume that the formation of a slip embryo is accompanied by the following free energy changes

$$+\gamma A \ldots\ldots [a]$$

$$-W^*V \ldots\ldots [b]$$

where γ is the interfacial energy per unit area of slip embryo, A is the area of the embryo, W^* is the strain energy per unit volume released by the formation of the embryo, and V is the volume throughout which the energy W^* has been released. The free energy change in the neighborhood of a circular slip embryo is therefore

$$\Delta F = \gamma A - W^* V$$

$$= \gamma (\pi r^2) - W^* \left(\frac{4}{3} \pi r^3\right) \ldots\ldots [1]$$

It is evident that the plot of ΔF versus r passes through a maximum as shown in Fig. 3, and that once a slip embryo exceeds the critical radius r^*, corresponding to this maximum, it becomes a slip nucleus and can continue to grow with monotonically decreasing free energy.

The theory of nucleation (7, 8) demonstrates that slip embryos

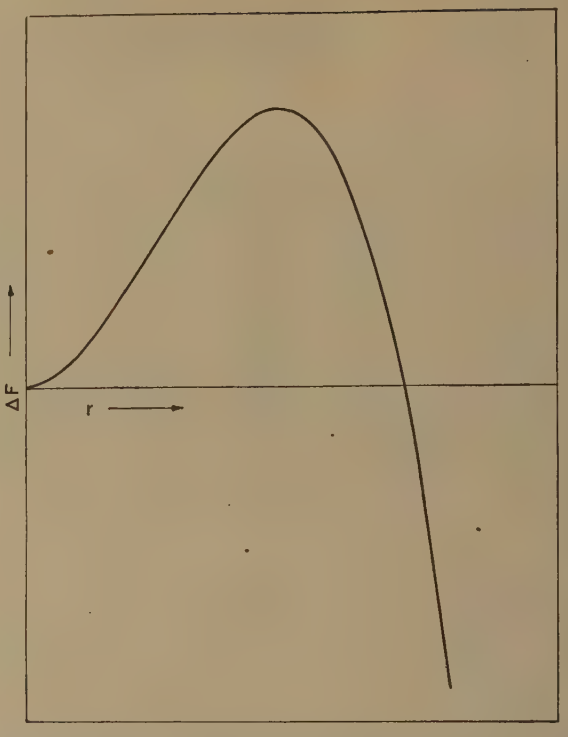


FIG. 3 FREE ENERGY CHANGE IN NEIGHBORHOOD OF A SLIP EMBRYO VERSUS RADIUS OF EMBRYO

change their sizes at finite rates, and that slip nuclei arise from slip embryos that have grown to critical size. Nucleation theory further states that for a given temperature and stress system, the rate of slip band nucleation in an ideal material will approach a steady-state value

$$\frac{dn}{dt} = Be^{-\Delta F^*/kT} \ldots\ldots [2]$$

where

$$\Delta F^* = \frac{\pi}{12} \frac{\gamma^3}{W^{*2}} \ldots\ldots [3]$$

is the value of ΔF corresponding to the critical radius r^*, and where B is a constant.

Plastic deformation carried out at a constant strain rate, i.e., a constant rate of nucleation of slip bands, will require

$$\frac{dn}{dt} = \text{const}$$

which at constant temperature requires

$$W^* = \text{const} \ldots\ldots [4]$$

In other words, in order to maintain a constant strain rate in an ideal material at a given temperature, it is necessary only to apply a constant value of W^*, the strain energy per unit volume released by slip.

Real crystals invariably contain defects which may either produce local stress concentrations or act as finite preformed slip embryos. Further, new defects are introduced by the deformation process itself. The rate of slip nucleation is sensitive to the size distribution of such defects. Since the distribution changes as deformation proceeds, the constant in Equation [4] will assume for real materials a characteristic value for each amount of deformation. In general

$$W^* = f \text{ (deformation)} \quad [5]$$

for plastic flow at a constant strain rate and temperature.

An "effective stress" σ may be defined proportional to the square root of W^*

$$\sigma = c\sqrt{W^*} \quad [6]$$

An "effective strain" ϵ may then be defined for an infinitesimal element of material according to the relationship

$$\int \sigma d\epsilon = W \quad [7]$$

where W is the work done per unit volume while deforming the element. It will be assumed that the effective strain

$$\epsilon = \int \frac{dW}{\sigma} \quad [8]$$

is a satisfactory measure of deformation. With these definitions Equation [5] can be written as

$$\sigma = f(\epsilon) \quad [9]$$

The plastic strains observed in crystals are in reality a summation of the elastic strains relaxed in the neighborhood of a very large number of slip bands. It follows that the plastic strain increments in a material are proportional to the difference between (*a*) the elastic strains before slip and (*b*) those after slip in the regions of stress relaxation near slip bands. Symbolically

$$\epsilon_i \sim e_i - e_i' \quad [10]$$

where ϵ_i is a plastic strain, e_i is the corresponding elastic strain before slip, and e_i' is the corresponding elastic strain in the slipped region after slip.

The plastic stress-strain relationships required for a generalized mathematical theory of plasticity have been derived, assuming the slip mechanism of plastic flow. The essential features of the development of the generalized theory are as follows:

1 The definition of an effective stress proportional to the square root of the strain energy per unit volume released by slip in the neighborhood of a slip band,

2 The definition of an effective strain such that the integral of effective stress with respect to effective strain gives the work done in deforming a unit volume of material,

3 The assumption that the effective strain (or the work done per unit deformed volume) is a satisfactory measure of deformation,

4 A demonstration that there exists a functional relationship between the effective strain and the effective stress required to maintain a constant strain rate at a constant temperature,

5 A demonstration that the plastic strain increments are proportional to the difference between (*a*) the elastic strains before slip and (*b*) the elastic strains in the slip region after slip.

It is of interest to apply the theory thus summarized to the plastic flow of single crystals and of isotropic polycrystals.

Flow of Single Crystals

In a single crystal the only stress relievable during slip is the resolved shear stress τ_* on the slip plane. The strain energy per unit volume relieved in the neighborhood of a slip band during slip is therefore

$$W^* = \tau_*^2/2G_*$$

where G_* is the appropriate elastic modulus. Defining

$$\sigma = \sqrt{2G_*}\sqrt{W^*} = \tau_* \quad [11]$$

proportional to the square root of W^*, the effective stress σ is identical with the resolved shear stress τ_*. The plastic strain increment is proportional to the elastic resolved shear strain, since this is the only strain that is relaxed by slip. The only plastic strain is, therefore, the resolved shear strain γ_*. The differential work required to deform plastically an element of material from resolved shear strain γ_* to $\gamma_* + d\gamma_*$ is

$$dW = \tau_* d\gamma_*$$

and the effective strain increment becomes

$$d\epsilon = \frac{dW}{\sigma} = \frac{\tau_* d\gamma_*}{\tau_*} = d\gamma_* \quad [12]$$

For monotonic plastic flow this can be integrated to give

$$\epsilon = \gamma_*$$

The relationship $\sigma = f(\epsilon)$ therefore becomes

$$\tau_* = f(\gamma_*) \quad [13]$$

for monotonic extension or compression of single crystals, giving the accepted resolved shear stress - shear strain relationship.

Flow of Isotropic Polycrystals

Each individual grain of an isotropic polycrystal must deform in a manner consistent with continuity of the entire matrix. Owing to interference from neighboring grains, generally it will be impossible for a grain to slip on a single set of slip planes. Therefore, the typical slip process will be that of simultaneous or alternate slip on a different set of planes in each of several contiguous regions (deformation bands) in the grain, the net displacements at the boundary of the grain conforming to the deformations of its neighbors (2). Slip on several different sets of slip planes allows more strain energy to be relaxed than does slip on a single set of planes. For isotropic polycrystals, therefore, it will be assumed that the average slip process relieves all shear stresses in the slip region, reducing the stress system to hydrostatic tension or compression.

Before slip, the energy per unit volume is

$$\left.\begin{aligned} W_1 &= \frac{1}{2}(\sigma_1 e_1 + \sigma_2 e_2 + \sigma_3 e_3) \\ &= \frac{1}{2E}[\sigma_1^2 + \sigma_2^2 + \sigma_3^2 - 2\nu(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1)] \end{aligned}\right\} \quad [14]$$

where σ_1, σ_2, σ_3 are principal stresses. After slip the principal stresses in the slip region are

$$\sigma_1' = \sigma_2' = \sigma_3' = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3)$$

since there is no volume change. The energy per unit volume remaining after slip is

$$\left.\begin{aligned} W_2 &= \frac{1}{2}(\sigma_1' e_1' + \sigma_2' e_2' + \sigma_3' e_3') \\ &= \frac{1-2\nu}{6E}(\sigma_1 + \sigma_2 + \sigma_3)^2 \end{aligned}\right\} \quad [15]$$

The energy per unit volume relievable by slip is therefore

$$\left.\begin{aligned} W^* &= W_1 - W_2 \\ &= \frac{1}{12G}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2] \end{aligned}\right\} \quad [16]$$

and the effective stress may be defined as

$$\left.\begin{aligned}\sigma &= \sqrt{6G}\,\sqrt{W^*} \\ &= \frac{\sqrt{2}}{2}\sqrt{(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2}\end{aligned}\right\} \quad \ldots\ldots [17]$$

The plastic strain increments are given by the relationships

$$\left.\begin{aligned} d\epsilon_1 &\sim \frac{\partial W_1}{\partial \sigma_1}-\frac{\partial W_2}{\partial \sigma_1'}=\frac{2(1+\nu)}{3E}\left[\sigma_1-\frac{1}{2}(\sigma_2+\sigma_3)\right] \\ d\epsilon_2 &\sim \frac{\partial W_1}{\partial \sigma_2}-\frac{\partial W_2}{\partial \sigma_2'}=\frac{2(1+\nu)}{3E}\left[\sigma_2-\frac{1}{2}(\sigma_3+\sigma_1)\right] \\ d\epsilon_3 &\sim \frac{\partial W_1}{\partial \sigma_3}-\frac{\partial W_2}{\partial \sigma_3'}=\frac{2(1+\nu)}{3E}\left[\sigma_3-\frac{1}{2}(\sigma_1+\sigma_2)\right] \\ d\gamma_{ij} &\sim \frac{\partial W_1}{\partial \tau_{ij}}-\frac{\partial W_2}{\partial \tau_{ij}'}=0\end{aligned}\right\} \quad \ldots\ldots [18]$$

The only plastic strains are therefore normal strains in the directions of the principal stresses

$$\left.\begin{aligned} d\epsilon_1 &= D\left[\sigma_1-\frac{1}{2}(\sigma_2+\sigma_3)\right] \\ d\epsilon_2 &= D\left[\sigma_2-\frac{1}{2}(\sigma_3+\sigma_1)\right] \\ d\epsilon_3 &= D\left[\sigma_3-\frac{1}{2}(\sigma_1+\sigma_2)\right]\end{aligned}\right\} \quad \ldots\ldots [19]$$

The differential work expended in deforming a plastic element of material is

$$dW = \sigma_1 d\epsilon_1 + \sigma_2 d\epsilon_2 + \sigma_3 d\epsilon_3$$

and the corresponding differential effective strain is

$$d\epsilon = \frac{dW}{\sigma} = \frac{1}{\sigma}(\sigma_1 d\epsilon_1 + \sigma_2 d\epsilon_2 + \sigma_3 d\epsilon_3) \quad \ldots\ldots [20]$$

which reduces to

$$d\epsilon = \sigma D$$

or

$$D = \frac{d\epsilon}{\sigma} \quad \ldots\ldots [21]$$

Noting that

$$d\epsilon_1 - d\epsilon_2 = \frac{3}{2} D(\sigma_1 - \sigma_2)$$

and so on, the expression for $d\epsilon$ becomes

$$d\epsilon = \frac{\sqrt{2}}{3}\sqrt{(d\epsilon_1-d\epsilon_2)^2+(d\epsilon_2-d\epsilon_3)^2+(d\epsilon_3-d\epsilon_1)^2} \quad \ldots [22]$$

The effective stress and effective strain for plastic deformation of an isotropic polycrystal have been shown to be

$$\left.\begin{aligned}\sigma &= \frac{\sqrt{2}}{2}\sqrt{(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2} \\ \epsilon &= \frac{\sqrt{2}}{3}\int\sqrt{(d\epsilon_1-d\epsilon_2)^2+(d\epsilon_2-d\epsilon_3)^2+(d\epsilon_3-d\epsilon_1)^2}\end{aligned}\right\} \quad \ldots\ldots [23]$$

The unique relationship $\sigma = f(\epsilon)$ is identical with that obtained from a tension test, $\sigma_x = f(\epsilon_x)$, as $\sigma = \sigma_x$ and $\epsilon = \epsilon_x$ for simple tension. The plastic strain increments are related to the stresses by the equations

$$\left.\begin{aligned} d\epsilon_1 &= \frac{d\epsilon}{\sigma}\left[\sigma_1-\frac{1}{2}(\sigma_2+\sigma_3)\right] \\ d\epsilon_2 &= \frac{d\epsilon}{\sigma}\left[\sigma_2-\frac{1}{2}(\sigma_3+\sigma_1)\right] \\ d\epsilon_3 &= \frac{d\epsilon}{\sigma}\left[\sigma_3-\frac{1}{2}(\sigma_1+\sigma_2)\right]\end{aligned}\right\} \quad \ldots\ldots [24]$$

The relationships summarized in Equations [23], [24] are those of the well-known distortion energy theory of plastic flow, which has received satisfactory experimental verification for isotropic polycrystals.

The success of the generalized theory of plasticity in reducing to the resolved shear stress-shear strain relationship for single crystals and to the distortion energy theory for isotropic polycrystals suggests that the generalized theory may prove of value for the analysis of plastic flow in anisotropic polycrystals.

Flow of Anisotropic Polycrystals

A particularly simple yet interesting and important type of polycrystalline anisotropy occurs when the material possesses an axis of symmetry in the sense that all rays at right angles to this axis are equivalent. Anisotropy of this type is found, for example, in rolled and extruded bars.

It will be recalled that according to the distortion energy theory of plastic flow, the differential plastic strains are linear functions of the stresses. In the present analysis attention will be directed to those cases of plastic flow where also the differential plastic strains are linear functions of the stresses. The problem is further simplified when the direction of the axis of symmetry is a principal stress direction.

With these limitations, taking the 1-axis coincident with the axis of symmetry and taking the 2- and 3-axes in the other principal stress directions, the principal stresses applied to an element of material are

$$\sigma_1,\ \sigma_2,\ \sigma_3$$

After slip, the stresses in the slipped region are linear functions of the applied stresses

$$\left.\begin{aligned}\sigma_1' &= \alpha_{11}\sigma_1+\alpha_{12}\sigma_2+\alpha_{13}\sigma_3 \\ \sigma_2' &= \alpha_{21}\sigma_1+\alpha_{22}\sigma_2+\alpha_{23}\sigma_3 \\ \sigma_3' &= \alpha_{31}\sigma_1+\alpha_{32}\sigma_2+\alpha_{33}\sigma_3 \\ \tau_{23}' &= \alpha_{41}\sigma_1+\alpha_{42}\sigma_2+\alpha_{43}\sigma_3 \\ \tau_{31}' &= \alpha_{51}\sigma_1+\alpha_{52}\sigma_2+\alpha_{53}\sigma_3 \\ \tau_{12}' &= \alpha_{61}\sigma_1+\alpha_{62}\sigma_2+\alpha_{63}\sigma_3\end{aligned}\right\} \quad \ldots\ldots [25]$$

since otherwise the differential plastic strains would not be linear functions of the applied stresses.

Constancy of volume in simple tension requires

$$\left.\begin{aligned}\alpha_{11}+\alpha_{21}+\alpha_{31} &= 1 \\ \alpha_{12}+\alpha_{22}+\alpha_{32} &= 1 \\ \alpha_{13}+\alpha_{23}+\alpha_{33} &= 1\end{aligned}\right\} \quad \ldots\ldots [26]$$

while the absence of slip under hydrostatic stress systems requires

$$\left.\begin{aligned}\alpha_{11}+\alpha_{12}+\alpha_{13} &= 1 \\ \alpha_{21}+\alpha_{22}+\alpha_{23} &= 1 \\ \alpha_{31}+\alpha_{32}+\alpha_{33} &= 1\end{aligned}\right\} \quad \ldots\ldots [27]$$

Assuming elastic isotropy, in anticipation of an application to polycrystalline aluminum, the strain energy before slip is

$$\left.\begin{aligned}W_1 &= \frac{1}{2}(\sigma_1 e_1+\sigma_2 e_2+\sigma_3 e_3) \\ &= \frac{1}{2E}[\sigma_1^2+\sigma_2^2+\sigma_3^2-2\nu(\sigma_1\sigma_2+\sigma_2\sigma_3+\sigma_3\sigma_1)]\end{aligned}\right\} \quad \ldots [28]$$

That in the slipped region after slip is

$$\left.\begin{aligned} W_2 &= \frac{1}{2}(\sigma_1'e_1' + \sigma_2'e_2' + \sigma_3'e_3') \\ &\quad + \frac{1}{2}(\tau_{23}'\gamma_{23}' + \tau_{31}'\gamma_{31}' + \tau_{12}'\gamma_{12}') \\ &= \frac{1}{2E}[\sigma_1'^2 + \sigma_2'^2 + \sigma_3'^2 \\ &\quad - 2\nu(\sigma_1'\sigma_2' + \sigma_2'\sigma_3' + \sigma_3'\sigma_1')] \\ &\quad + \frac{1}{2G}(\tau_{23}'^2 + \tau_{31}'^2 + \tau_{12}'^2) \end{aligned}\right\} \quad [29]$$

The corresponding plastic strain increments are

$$\left.\begin{aligned} d\epsilon_1 &\sim \frac{\partial W_1}{\partial \sigma_1} - \frac{\partial W_2}{\partial \sigma_1'} = \frac{1+\nu}{E}[(1-\alpha_{11})\sigma_1 - \alpha_{12}\sigma_2 - \alpha_{13}\sigma_3] \\ d\epsilon_2 &\sim \frac{\partial W_1}{\partial \sigma_2} - \frac{\partial W_2}{\partial \sigma_2'} = \frac{1+\nu}{E}[-\alpha_{21}\sigma_1 + (1-\alpha_{22})\sigma_2 - \alpha_{23}\sigma_3] \\ d\epsilon_3 &\sim \frac{\partial W_1}{\partial \sigma_3} - \frac{\partial W_2}{\partial \sigma_3'} = \frac{1+\nu}{E}[-\alpha_{31}\sigma_1 - \alpha_{32}\sigma_2 + (1-\alpha_{33})\sigma_3] \\ d\gamma_{23} &\sim \frac{\partial W_1}{\partial \tau_{23}} - \frac{\partial W_2}{\partial \tau_{23}'} = \frac{1}{G}[\alpha_{41}\sigma_1 + \alpha_{42}\sigma_2 + \alpha_{43}\sigma_3] \\ d\gamma_{31} &\sim \frac{\partial W_1}{\partial \tau_{31}} - \frac{\partial W_2}{\partial \tau_{31}'} = \frac{1}{G}[\alpha_{51}\sigma_1 + \alpha_{52}\sigma_2 + \alpha_{53}\sigma_3] \\ d\gamma_{12} &\sim \frac{\partial W_1}{\partial \tau_{12}} - \frac{\partial W_2}{\partial \tau_{12}'} = \frac{1}{G}[\alpha_{61}\sigma_1 + \sigma_{62}\sigma_2 + \alpha_{63}\sigma_3] \end{aligned}\right\} \quad [30]$$

Symmetry of the deformations accompanying plastic flow in simple tension in the principal directions requires that the plastic-strain increments $d\gamma_{23}$, $d\gamma_{31}$, $d\gamma_{12}$ shall all be zero. This is possible only when

$$\alpha_{41} = \alpha_{42} = \alpha_{43} = \alpha_{51} = \alpha_{52} = \alpha_{53} = \alpha_{61} = \alpha_{62} = \alpha_{63} = 0 \quad [31]$$

Symmetry further requires

$$\left.\begin{aligned} \alpha_{12} &= \alpha_{13} \quad \alpha_{23} = \alpha_{32} \\ \alpha_{21} &= \alpha_{31} \quad \alpha_{22} = \alpha_{33} \end{aligned}\right\} \quad [32]$$

The expressions for the plastic-strain increments, therefore, simplify to

$$\left.\begin{aligned} d\epsilon_1 &= D_1\left[(1-\alpha_{11})\sigma_1 - \frac{1}{2}(1-\alpha_{11})\sigma_2 - \frac{1}{2}(1-\alpha_{11})\sigma_3\right] \\ d\epsilon_2 &= D_1\left[-\frac{1}{2}(1-\alpha_{11})\sigma_1 + (1-\alpha_{22})\sigma_2 - \frac{1}{2}(1+\alpha_{11}-2\alpha_{22})\sigma_3\right] \\ d\epsilon_3 &= D_1\left[-\frac{1}{2}(1-\alpha_{11})\sigma_1 - \frac{1}{2}(1+\alpha_{11}-2\alpha_{22})\sigma_2 + (1-\alpha_{22})\sigma_3\right] \end{aligned}\right\} \quad [33]$$

where the substitutions

$$\left.\begin{aligned} \alpha_{11} &= \alpha_{11} & \alpha_{12} &= \frac{1}{2}(1-\alpha_{11}) \\ \alpha_{21} &= \frac{1}{2}(1-\alpha_{11}) & \alpha_{22} &= \alpha_{22} \\ \alpha_{31} &= \frac{1}{2}(1-\alpha_{11}) & \alpha_{32} &= \frac{1}{2}(1+\alpha_{11}-2\alpha_{22}) \\ & & \alpha_{13} &= \frac{1}{2}(1-\alpha_{11}) \\ & & \alpha_{23} &= \frac{1}{2}(1+\alpha_{11}-2\alpha_{22}) \\ & & \alpha_{33} &= \alpha_{22} \end{aligned}\right\} \quad [34]$$

have been made.

The strain energy per unit volume released by slip may be computed as $W^* = W_1 - W_2$ which reduces to

$$\frac{4EW^*}{(1+3\alpha_{11})(1-\alpha_{11})(1+\nu)} = \frac{1}{2}[(\sigma_1-\sigma_2)^2 + (\sigma_2-\sigma_3)^2 + (\sigma_3-\sigma_1)^2] + \frac{m}{2}(\sigma_2-\sigma_3)^2 \quad [35]$$

where

$$m = \frac{4(\alpha_{22}-\alpha_{11})(1-\alpha_{11}-2\alpha_{22})}{(1+3\alpha_{11})(1-\alpha_{11})} \quad [36]$$

The effective stress may be defined as

$$\left.\begin{aligned} \sigma &= \sqrt{4E/(1+3\alpha_{11})(1-\alpha_{11})(1+\nu)}\ \sqrt{W^*} \\ &= \frac{\sqrt{2}}{2}\sqrt{(\sigma_1-\sigma_2)^2 + (1+m)(\sigma_2-\sigma_3)^2 + (\sigma_3-\sigma_1)^2} \end{aligned}\right\} \quad [37]$$

Setting $D = (1-\alpha_{11})D_1$ and defining

$$\beta = \frac{1+\alpha_{11}-2\alpha_{22}}{1-\alpha_{11}} \quad [38]$$

the values of $d\epsilon_1$, $d\epsilon_2$, $d\epsilon_3$ may be rewritten

$$\left.\begin{aligned} d\epsilon_1 &= D\left[\sigma_1 - \frac{1}{2}\sigma_2 - \frac{1}{2}\sigma_3\right] \\ d\epsilon_2 &= D\left[-\frac{1}{2}\sigma_1 + \frac{1+\beta}{2}\sigma_2 - \frac{\beta}{2}\sigma_3\right] \\ d\epsilon_3 &= D\left[-\frac{1}{2}\sigma_1 - \frac{\beta}{2}\sigma_2 + \frac{1+\beta}{2}\sigma_3\right] \end{aligned}\right\} \quad [39]$$

The effective strain increment

$$d\epsilon = \frac{dW}{\sigma} = \frac{1}{\sigma}(\sigma_1 d\epsilon_1 + \sigma_2 d\epsilon_2 + \sigma_3 d\epsilon_3) \quad [40]$$

can be expressed in terms of $d\epsilon_1$, $d\epsilon_2$, $d\epsilon_3$, m, β by making use of the relationships in Equations [37], [39], and [40]

$$d\epsilon = \frac{\beta d\epsilon_1^2 + d\epsilon_2^2 + d\epsilon_3^2}{\frac{\sqrt{2}}{2}\sqrt{(\beta d\epsilon_1 - d\epsilon_2)^2 + (1+m)(d\epsilon_2 - d\epsilon_3)^2 + (d\epsilon_3 - \beta d\epsilon_1)^2}} \quad [41]$$

Equations [36], [37], [38] [39], [40] [41] summarize the plastic stress-strain relationships of a typical anisotropic polycrystal of the type under consideration.

The quantities m and β may be thought of as plastic anisotropy parameters having the values $m = 0$ and $\beta = 1$ for isotropic media. Both m and β can be determined experimentally from two tension tests, one in the 1-direction (the direction of the axis of symmetry) and one in the 2- or 3-direction. Consider first a tension test where the axis of the specimen is in the 2-direction of the material. The plastic strain increments during this test are

$$\left.\begin{aligned} d\epsilon_1 &= D\left[-\frac{1}{2}\sigma_2\right] \\ d\epsilon_2 &= D\left[\frac{1+\beta}{2}\sigma_2\right] \\ d\epsilon_3 &= D\left[-\frac{\beta}{2}\sigma_2\right] \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots [39a]$$

and the ratio of lateral strains is

$$d\epsilon_3/d\epsilon_1 = \beta \quad \ldots\ldots\ldots\ldots\ldots [42]$$

giving directly the value of β from the ellipticity of the cross section.

For the same tension test with $\sigma_2 = \sigma_2$, $\sigma_1 = \sigma_3 = 0$ as before, the effective stress is

$$\sigma = \frac{\sqrt{2}}{2}\sqrt{\sigma_2^2 + (1+m)\sigma_2^2} = \sigma_2\sqrt{1+\frac{m}{2}} \ldots [37a]$$

The effective strain is

$$d\epsilon = \frac{\sigma_2 d\epsilon_2}{\sigma} = \frac{\sigma_2 d\epsilon_2}{\sigma_2\sqrt{1+\frac{m}{2}}} = d\epsilon_2\Big/\sqrt{1+\frac{m}{2}} \ldots [41a]$$

or for monotonic extension

$$\epsilon = \epsilon_2\Big/\sqrt{1+\frac{m}{2}}$$

For a tension test in the 1-direction the effective stress reduces to

$$\sigma = \frac{\sqrt{2}}{2}\sqrt{\sigma_1^2 + \sigma_1^2} = \sigma_1 \quad \ldots\ldots\ldots\ldots [37b]$$

and the corresponding effective strain is

$$d\epsilon = \sigma_1 d\epsilon_1/\sigma = d\epsilon_1 \quad \ldots\ldots\ldots\ldots\ldots [41b]$$

or for monotonic extension

$$\epsilon = \epsilon_1$$

Since the effective stress-effective strain relationship is invariant, it will be possible to find a value of m such that the plot of

$$\sigma = \sigma_2\sqrt{1+\frac{m}{2}} \quad \text{versus} \quad \epsilon = \epsilon_2\Big/\sqrt{1+\frac{m}{2}}$$

coincides with the plot of

$$\sigma = \sigma_1 \text{ versus } \epsilon = \epsilon_1$$

Having determined the values of m and β for a given plastically anisotropic material, the behavior of the material under arbitrary stress systems (with a principal stress in the direction of the axis of symmetry) can be predicted from the theory.

Application of Anisotropic-Flow Theory

Plastic flow corresponding to various systems of combined stresses is frequently studied by subjecting tubes to internal pressure and end load. It is of interest to examine the behavior of such a tube cut from a plastically anisotropic material, the axis of the tube coinciding with the axis of anisotropic symmetry.

The recent analysis of large plastic deformations in isotropic thick-walled tubes (9) can be extended readily to the deformations of anisotropic tubes. Appropriate modification gives the following stresses and plastic strains in an anisotropic closed-end thick-walled tube subjected to internal pressure, where the 1-direction is now written as the z-direction

$$\left.\begin{aligned} \sigma_r &= -\frac{2}{\sqrt{3}}\left(1+\frac{2m}{3}\right)^{-\frac{1}{2}}\int_r^{R_1}\sigma e^{-2\epsilon_t}\,\frac{dr}{r} \\ \sigma_t &= \sigma_r + \frac{2\sigma}{\sqrt{3}}\left(1+\frac{2m}{3}\right)^{-\frac{1}{2}} \\ \sigma_z &= \frac{1}{2}(\sigma_r + \sigma_t) \\ \epsilon_t &= \frac{1}{2}\log_e\left[1+\frac{R_1^2}{r^2}(e^{2\epsilon_{t1}}-1)\right] \\ \epsilon_r &= -\epsilon_t \qquad \epsilon_z = 0 \\ \epsilon &= \frac{2\epsilon_t}{\sqrt{3}}\left(1+\frac{2m}{3}\right)^{-\frac{1}{2}} \end{aligned}\right\} \quad \ldots [43]$$

Here σ_r, σ_t, σ_z (ϵ_r, ϵ_t, ϵ_z) are the radial, tangential, and axial stresses (strains), respectively; r is the radial distance from the axis of the tube to an arbitrary element; R_1 is the value of r at the external surface of the tube; and ϵ_{t1} is the value of ϵ_t at $r = R_1$. It is of interest to note that when $\epsilon_z = 0$ only one of the two anisotropy parameters (m, β) appears in the expressions for the stresses and strains.

The predictions of the anisotropic-flow theory have been investigated using an extruded aluminum bar, 10 ft long and one-half ft diam, as the source of material. A number of tension tests indicated that this material was plastically anisotropic, with the axis of the extruded bar being an axis of symmetry in the sense that all directions at right angles to it were equivalent. Calling the direction of the axis of the bar (and of anisotropic symmetry) the axial direction, it was found that tension specimens having axes coincident with the axial direction remained circular in cross section during extension, while tension specimens having axes coincident with any lateral direction at right angles to the axial direction became elliptical in cross section during extension.

The value of $\beta = \epsilon_3/\epsilon_1$ corresponding to a tension test of a specimen with axis in the 2-direction was determined for a number of specimens cut in various lateral directions. The mean value of β so found was

$$\beta = 1.65 \quad \ldots\ldots\ldots\ldots\ldots [42a]$$

indicating a moderate degree of anisotropy.

The value of m was determined so that the mean curve of

$$\sigma_z \text{ versus } \epsilon_z$$

for several axial tension specimens coincided with the mean curve of

$$\sigma_t\sqrt{1+\frac{m}{2}} \quad \text{versus} \quad \epsilon_t\Big/\sqrt{1+\frac{m}{2}}$$

for several lateral tension specimens. Fig. 4 shows that the mean curves

$$\sigma_z \quad \text{versus} \quad \epsilon_z$$

and

$$1.07\ \sigma_t \text{ versus } \epsilon_t/1.07$$

do in fact coincide within the limit of accuracy of the measurements. The value of m corresponding to $(1 + m/2)^{1/2} = 1.07$ is

$$m = 0.29$$

Having the value of m and the relationship in Fig. 4 between the effective stress σ and the effective strain ϵ, it is possible to calculate by means of Equations [43] a relationship between the

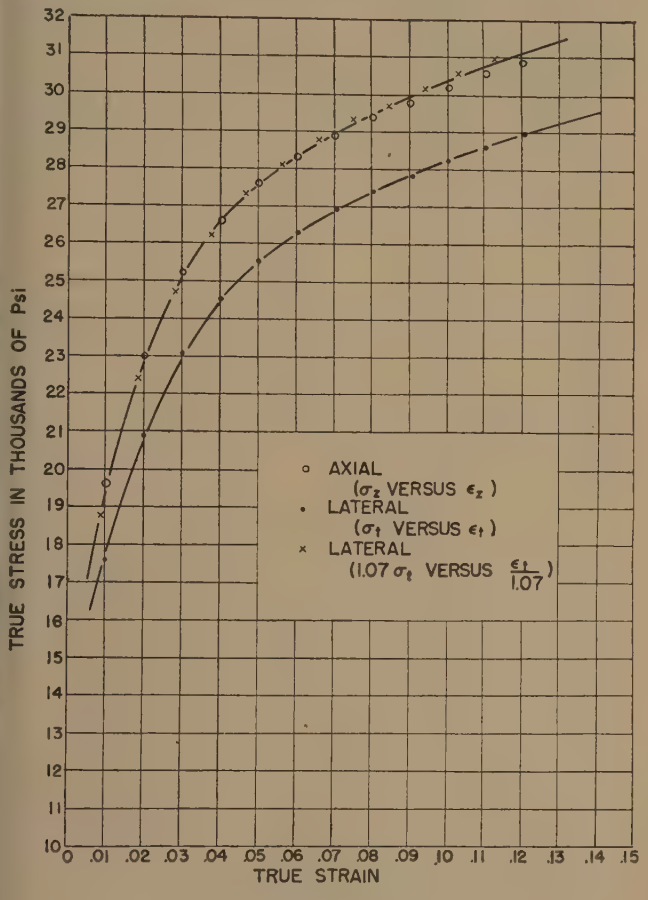


FIG. 4 MEAN STRESS-STRAIN CURVES FOR ALUMINUM

internal pressure and the corresponding external tangential strain for the expansion of a closed-end thick-walled tube of the material in question. Such a calculation, involving numerical solution of Equations [43], has been carried out for a tube with wall ratio 2. The theoretical pressure-expansion curve is given as curve B in Fig. 5. Curve A was calculated using the distortion energy theory based on the mean tangential stress-strain relationship. Had curve A been based on the mean axial stress-strain relationship, it would have fallen still higher than curve B. Curve C was determined experimentally from a tube with an initial external diameter of 5 in. and an initial internal diameter of 2.5 in. The length of the tube including the threaded ends which were fitted with steel caps was 36 in., while the uniform section was 24 in. long.

It is evident that the prediction of the anisotropic flow theory is significantly better than that of the distortion energy theory, differing from experiment by only about 4 per cent compared to a minimum of about 9 per cent for the predictions of the distortion energy theory.

Although it did not enter the calculation of the stresses and de-

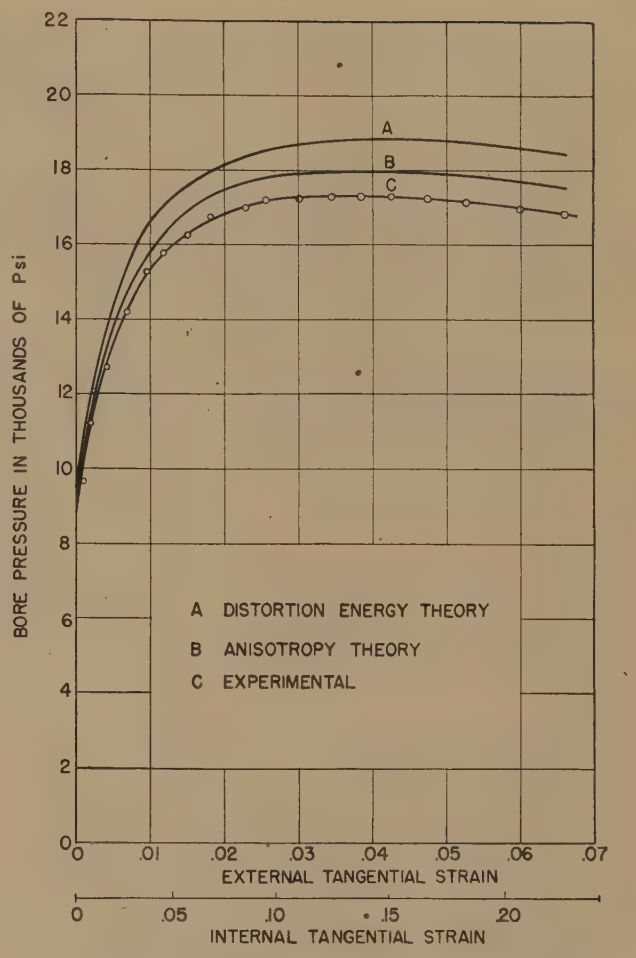


FIG. 5 PRESSURE-EXPANSION CURVES FOR ALUMINUM TUBE

formations in a closed-end thick-walled tube, the value of β will be important whenever the strain in the direction of the axis of plastic symmetry is different from zero. For example, a combined stress test where $\sigma_1 = \sigma_2$, $\sigma_3 = 0$ leads to an effective stress

$$\sigma = \sigma_1 \sqrt{1 + \frac{m}{2}}$$

and an effective strain

$$\epsilon = (1 + \beta)\epsilon_1 \Big/ \sqrt{1 + \frac{m}{2}}$$

Many investigations of the effects of combined stresses on the plastic flow of metals have been made by subjecting tubes, machined from rolled or extruded bars, to internal pressure and end load. Since plastic anisotropy is to be expected in these materials, many of the discrepancies between experiment and the predictions of the distortion energy theory may be attributed to anisotropy. The results summarized in Fig. 5 indicate, for example, that the pressure-expansion curve for a closed-end tube can fall below that predicted by the distortion energy theory based on the lower of the two stress-strain curves for an anisotropic material.

ACKNOWLEDGMENT

The experiments described in this paper were sponsored by the Bureau of Ordnance, Navy Department (Contract NOrd-9107) and were carried out at the Massachusetts Institute of Tech-

nology under the direction of Prof. C. W. MacGregor. The analysis of anisotropic plastic flow was contained in part in the thesis presented by the author in partial fulfillment of the requirements for the degree of Sc.D. at the Massachusetts Institute of Technology.

The author would like to express his thanks to Prof. C. W. MacGregor and Prof. L. F. Coffin, Jr., for their interest and encouragement while he was at the Massachusetts Institute of Technology, and to Dr. J. H. Hollomon and the Physical Metallurgy Group at the General Electric Company for making it possible for him to continue the development of the theory.

BIBLIOGRAPHY

1 "The Distortion of Metal Crystals," by C. F. Elam, Oxford University Press, New York, N. Y., 1935.

2 "Structure of Metals," by C. S. Barrett, McGraw-Hill Book Company, Inc., New York, N. Y., 1943.

3 "Mechanik der festen Körper im plastisch-deformablen Zustand," by R. von Mises, *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen*, 1913, 582–592.

4 "Zur Theorie plastischer Deformationen, etc.," by H. Hencky, *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 4, 1924, p. 323.

"Uber das Wesen der plastischen Verformung," by H. Hencky, *Zeitschrift des Vereines dentscher Ingenieure*, vol. 69, 1925, p. 695.

5 "The Physics of Crystals," by A. F. Joffe, McGraw-Hill Book Company, Inc., New York, N. Y., 1928.

6 "The Latent Energy Remaining in a Metal After Cold Working," by G. I. Taylor and H. Quinney, Proceedings of the Royal Society of London, England, vol. A143, 1934, p. 307.

7 "Nucleation of Slip Bands," by J. G. Leschen, R. P. Carreker, and J. H. Hollomon, Trans AIME, Metals Technology, Technical Publication No. 2476, September, 1948.

8 "Nucleation," by J. C. Fisher, J. H. Hollomon, and D. Turnbull, *Journal of Applied Physics*, vol. 19, 1948, p. 775.

9 "The Plastic Flow of Thick-Walled Tubes With Large Strains," by C. W. MacGregor, L. F. Coffin, Jr., and J. C. Fisher, *Journal of Applied Physics*, vol. 19, 1948, p. 291.

The Laminar-Film Hypothesis

By BENJAMIN MILLER,[2] OZONE PARK, N. Y.

The laminar-film hypothesis which plays so prominent a part in current heat-transfer literature grew out of a formula relating the pressure gradient along a tube through which a fluid is passing to the rate of heat transfer between the tube and the fluid. Perry derived this formula by combining Péclet's idea that heat transferred between a body of fluid in violent motion and a bounding solid must flow by conduction through a fluid film at the boundary, with Reynolds analogy between heat convection and momentum convection. Prandtl and later Taylor identified the surface film, invented to explain heat-transfer findings, with a layer of fluid assumed in fluid mechanics to move in laminar flow along the tube wall while the body of fluid forms a turbulently moving core, and they modified Perry's formula accordingly. Stanton pointed out that the film thickness could be calculated by the modified formula from heat-transfer measurements, and that proof by velocity measurements that there was indeed a laminar film having the calculated thickness would substantiate the hypothesis; but his experiments showed only that there was a region characterized by velocity gradients intermediate between those in the central part of the tube and those calculated for laminar flow. Eagle and Ferguson offered a formula having a term for the intermediate zone as well as one for the laminar film and one for the turbulent core, and calculated the numerical constants from their heat-transfer measurements. Von Kármán adopted the three-zone scheme, but obtained his numerical constants from a velocity distribution published by Nikuradse. If the velocity distribution used by von Kármán had actually been determined experimentally by Nikuradse, the laminar-film hypothesis would be strongly supported, but Nikuradse's experimental results were inconsistent with the laminar-film hypothesis; having more faith in the hypothesis than in his experimental results, Nikuradse adjusted the results to make them seem to be in accord with the hypothesis, and published the adjusted results as well as the unadjusted, without explaining the discrepancy.

THE concept of a surface film, which resists the transfer of heat between a solid and the fluid in contact with it, is one of the most venerable in heat-transfer literature. Péclet (1)[3] invented it more than a hundred years ago to explain some astonishing results he obtained in experiments intended to determine the thermal conductivities of the metals. These experiments, begun in 1841, made use of a hollow cylinder fitted with a stirrer, and well insulated externally except for a detachable metal bottom, separated from the rest of the cylinder by a ring of cork. Steam condensed on the underside of the metal bottom to heat water contained in the cylinder, and observations of the temperature of the water at timed intervals were used to calculate the heat-transfer rate.

At first, Péclet was astonished to find that, with the quantity of water and its initial temperature maintained the same in a series of experiments, the time taken for the temperature to rise 1 deg was affected not at all by the kind of metal of which the bottom was made, nor by its thickness, although the time did decrease as the violence of agitation increased. He hypothesized that the metal disk was sandwiched between two layers of water whose thermal resistance was so much greater than that of the metal disk that any effect of variations in the latter was entirely obscured. Following this reasoning, he set up other apparatus in which a thick disk of poorly conducting lead was placed between two bodies of water whose temperatures differed by only a few degrees, and arranged for both sides of the disk to be scrubbed continuously during the run. When in two runs the passage of a certain quantity of heat took 500 sec with a disk 20 mm thick and 380 sec with one 15 mm thick, everything else being equal, Péclet was satisfied that the surface film was no longer interfering seriously, and that he was measuring the thermal conductivity of the metal itself.

Péclet's work was studied by that remarkable engineer Osborne Reynolds, who criticized it as being entirely empirical. In 1874, Reynolds (2) suggested that the design of steam boilers and other heat exchangers might be based rationally on the kinetic theory, which was then being developed by Maxwell and others. It followed from this theory that the rate at which heat can be taken from a surface by a fluid would be proportional to the temperature difference between fluid and surface, and to the rate at which molecules can pass back and forth between the surface and the fluid; for definiteness, between a line on the surface and a parallel line within the fluid. The rate of molecule passage depends on two things, namely, the natural movement of the molecules which exists when the fluid seems to be at rest, and the eddies which are caused by visible motion, and which continually bring fresh particles to the surface. Assuming that the eddy contribution is proportional to the velocity with which the fluid moves past the surface, and to the density of the fluid to allow for the effect of pressure if the fluid be a gas, Reynolds predicted that for heat transfer between a tube and a fluid flowing through it the law would be found to be

$$q_0 = A\,(t_w - t_f) + BV\rho\,(t_w - t_f) \quad [1]$$

where

q_0 = heat transferred from unit surface of wall to fluid in unit time
t_w = temperature of wall
t_f = temperature of fluid
V = flow velocity, volume of flow per unit time divided by cross-sectional area of tube
ρ = density of fluid
A, B = constants for each fluid

One of the formulas for flow resistance then in vogue was

$$\tau_0 = A'V + B'V^2\rho \quad [2]$$

where

τ_0 = resistance to flow per unit surface of wall
A', B' = constants for each fluid

Reynolds gave Equation [2] as the law of flow resistance, and

---

[1] Contributed by the Heat Transfer Division and presented at the Annual Meeting, Atlantic City, N. J., Dec. 1–5, 1947, of The American Society of Mechanical Engineers. Adapted from a lecture entitled, "Exploding a Heat Transfer Myth," delivered to the Nepa Heat Transfer Symposium at Oak Ridge, Tenn., December, 1947.

[2] Consulting Engineer.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

said that there were reasons to believe that A and B would be found to be proportional to A' and B'. This is the first form of the Reynolds analogy between heat transfer and resistance to flow.

In the 1874 paper (2) Reynolds mentioned some preliminary experiments he had made to test his prediction, and spoke of further experiments he hoped to make. It was not until 1897 (3) that further developments were reported. In that year there was published a paper by Stanton describing experiments on heat transfer to water flowing in tubes; the experiments had been made in Reynolds' laboratory in 1895 and 1896. The paper included a section on theory which Stanton attributed to Reynolds; this was supposedly based on the same ideas as those in the 1874 paper. In the years that had intervened, Reynolds had done the classic work on viscous and turbulent flow, and the transition, and he no longer accepted Equation [2] as the resistance law, but instead he had verified Poiseuile's law for viscous flow, and for turbulent flow he had put forth the Reynolds index law, which is

$$-\frac{dp}{dx} = A_2 \left(\frac{DV\rho}{\mu}\right)^n \frac{\mu^2}{\rho D^3} \quad [3]$$

where

p = pressure on fluid
x = distance along tube in flow direction
$\frac{dp}{dx}$ = rate of increase of pressure with distance along tube in flow direction
D = diameter of tube
μ = viscosity of fluid
A_2, n = constants

Reynolds reasoned that every molecule of the fluid which touched the wall would come to equilibrium with the wall, so that the change in temperature of a particular mass of fluid while passing through an elementary length of the tube would be a fraction of the difference in temperature between tube wall and entering fluid, that fraction being equal to the fraction of the molecules which touched the wall during the passage. The increase in temperature of the fluid while passing through the differential length of tube dx being designated dt, the fraction of the molecules which touched the wall during the passage is $(dt)/(t_w - t_f)$. Now every molecule which touches the wall loses its velocity with respect to the wall, thus giving up to the wall its momentum. The velocity of the fluid leaving the elementary length is not less than it was when entering, because the momentum is restored by the decrease in pressure. The rate at which momentum is regained by the fluid in the length dx is

$$-\frac{\pi D^2\, dp}{4}$$

The total mass of the fluid passing through the length dx in unit time is the product of its cross section $(\pi D^2)/4$, by its density ρ, by its velocity V, or

$$\frac{\rho\pi D^2}{4}V$$

and the total momentum of that mass is accordingly

$$\frac{\rho\pi D^2}{4}V^2$$

Then the ratio of the number of molecules which touch the wall while passing through the length dx to the total number which pass through is the ratio of the momentum gained to the total momentum, or

$$-\frac{dp}{\rho V^2}$$

Reynolds therefore equated the fraction obtained by considering temperature to the fraction obtained by considering velocity, and obtained

$$\frac{dt}{t_w - t_f} = -\frac{dp}{\rho V^2} \quad [4]$$

Equation [4] is the second form of the Reynolds analogy. Combining Equations [3] and [4]

$$\frac{dt}{dx} = A_2 \left(\frac{DV\rho}{\mu}\right)^{n-2} \frac{t_w - t_f}{D} \quad [5]$$

Reynolds pointed out that Equation [5] was derived on the assumption that the conductivity of the fluid, as compared to its viscosity, did not take part; but since the heat must pass by conduction from wall to fluid, he predicted that it would be necessary to introduce some function of the ratio[4] of the conductivity to the viscosity, which would have to be determined by experiment.

Stanton compared his experimental results with Equation [5], and found that the rate of heat transfer was considerably less than the formula predicted; he ascribed the discrepancy to the conductivity effect.

Since

$$q_0 = \frac{DV\rho C_p}{4}\frac{dt}{dx} \quad [6]$$

where

C_p = heat capacity of fluid

and

$$\tau_0 = -\frac{D}{4}\frac{dp}{dx} \quad [7]$$

Equation [4] is equivalent to

$$q_0 = \frac{\tau_0 C_p (t_w - t_f)}{V} \quad [8]$$

However, if the heat must first pass by conduction through a surface film, Equation [8] overestimates q_0, and the relation should be

$$q_0 = \frac{\tau_0 C_p (t_b - t_f)}{V} \quad [9]$$

where

t_b = temperature at inner boundary of film

To evaluate t_b we have

$$q_0 = \frac{(t_w - t_b)\, k}{b} \quad [10]$$

[4] It should be noted that water is the only fluid mentioned by Reynolds, and Equations [3], [4], and [5] have been changed from the forms given in the paper by inserting density in appropriate places to achieve dimensional homogeneity, making them applicable to fluids whose density is not unity. The "ratio of the conductivity to the viscosity" similarly must be changed to the "ratio of the conductivity to the product of the viscosity and the heat capacity" to make the statement applicable to fluids whose heat capacity is not unity.

where

b = thickness of film
k = thermal conductivity of fluid

Combining Equations [9] and [10]

$$q_0 = \frac{\tau_0 C_p (t_w - t_f)}{V} \cdot \frac{1}{1 + \frac{\tau_0 C_p b}{Vk}} \quad \ldots\ldots [11]$$

Equation [11] and its derivation are due to Perry (4), who said he was led to it by considering Stanton's paper.

It is implied in the derivation of Equation [11] that the fluid in the surface film does not move along the tube. However, Prandtl (5) pointed out that the fluid must move along the tube, except at the wall, and that the flow in the film must be viscous.[5] A molecule which touched the inner boundary of the film would not have its velocity reduced to zero, but only to the velocity of the fluid flowing at that boundary; this velocity may be designated U. Therefore the loss of momentum suffered by such a molecule would be proportional to $V - U$, not to V. Carrying through the derivation with this change, instead of Equation [9] one obtains

$$q_0 = \frac{\tau_0 C_p (t_b - t_f)}{V - U} \quad \ldots\ldots [12]$$

Actually, one should write τ_b rather than τ_0, to indicate the resistance to flow of the core per unit surface of boundary; if the thickness of the film is very small, the difference between τ_0 and τ_b is also very small, as can be shown by considering the force on the cross section of the whole tube and the force on the cross section of the core.

Assuming that the viscous film is thin enough so that the curvature can be neglected, the velocity will decrease linearly from the boundary to the wall, and the thickness of the film follows from the definition of viscosity

$$b = \frac{\mu U}{\tau_0} \quad \ldots\ldots [13]$$

Eliminating b by combining Equations [10] and [13], and then combining the result with Equation [12] to eliminate t_b

$$q_0 = \frac{\tau_0 C_p (t_w - t_f)}{V} \cdot \frac{1}{1 + \frac{U}{V}\left(\frac{C_p \mu}{k} - 1\right)} \quad \ldots\ldots [14]$$

Equation [14] reduces to Equation [8] if $C_p \mu/k$ is equal to unity; this group is called the Prandtl number, and will be designated N_{Pr}.

Equation [14] was independently derived by Taylor (6) a few years later. He decided on theoretical grounds that the ratio of U to V should be about 0.38 although the only measurements available to him indicated a much higher value.

Discussing Taylor's paper, Stanton (7) pointed out that measurements on heat transfer to water could be used to calculate the ratio, and he found it to be in the neighborhood of 0.3. He also said: "The simplest check on the accuracy of the theoretical conclusions would be to explore the region in the neighborhood of the surface by means of a fine Pitot tube and determine the value of U (the velocity at the outer surface of the layer of fluid which is undisturbed by the eddying motion and through which the heat is transmitted by conductivity) corresponding with a given value of V (the mean velocity of flow over the surface). During the last few years, several attempts to solve the problem by this method have been made at the National Physical Laboratory, and the work is still in progress. Up to the present, observations have only been made in the comparatively wide region over which a state of motion exists, which is intermediate between turbulent and laminar flow, and pure laminar flow has not been observed."[6]

It seems likely that by "pure laminar flow," or "laminar flow," Stanton meant flow such that the velocity distribution is given by

$$u = \frac{1}{2} \frac{\tau_0}{\mu} \left(\frac{r_0^2 - r^2}{r_0}\right) \quad \ldots\ldots [15]$$

where

u = velocity at any point
r = distance from tube axis to point
r_0 = distance from tube axis to wall, or $D/2$

while by "turbulent flow" he meant flow such that the velocity is substantially independent of distance from the axis. If this be so, a state of motion intermediate between laminar and turbulent flow would be one characterized by velocity gradients less than given by Equation [15], but still considerable. The present author prefers to use "turbulent" to indicate velocity gradients less than given by Equation [15], and to characterize flow as slightly, moderately, or highly turbulent, depending upon the difference between the actual velocity gradient and that given by Equation [15].

In 1920 Stanton, Marshall, and Bryant (8) described their measurements with fine Pitot tubes; they did not succeed in measuring velocities close to the wall, even when the flow was laminar throughout the tube.

In 1928 Prandtl (9) returned to the problem, and proposed a very ingenious scheme for estimating the thickness of the viscous film, based on an assumed velocity distribution developed from the Reynolds index law by using the values of the constants in the index law assigned by Blasius (10). The development of the assumed velocity distribution had been given in a previous paper by von Kármán (11) on dimensional grounds; the present author has found it easier to explain it as follows:

It is assumed that, outside the viscous film, the velocity increases with some power of the distance from the wall, so that

$$u = a y^m \quad \ldots\ldots [16]$$

where y = distance from wall, and a and m are constants. Then through any elementary ring of thickness dr the flux, which is $2 \pi r u\, dr$, will be $2 \pi a r (r_0 - r)^m\, dr$. Integrating from axis to pipe wall, and dividing the result by the cross-sectional area of the tube gives the average velocity, or

$$V = \frac{2\pi a \int_0^{r_0} r (r_0 - r)^m\, dr}{\pi r^2} \quad \ldots\ldots [17]$$

[5] Prandtl said: "In the case of turbulent flow through a tube convection is secondary to viscosity in the immediate vicinity of the wall, but plays the principal role in the central region. Between the two realms is a zone in which viscosity and convection are equally important; it is certainly wrong to say that the thickness of this zone is zero, that only a boundary line divides viscosity's realm from that in which convection is supreme, but it is only on this assumption that the calculation can be carried forward." Later authors have overlooked this passage.

[6] The crux of the matter is in this quotation from Stanton. The laminar-film hypothesis arises from heat-transfer considerations. Evidence from heat-transfer experiments may be used to support the hypothesis, but that evidence may be interpreted in various ways. Only velocity measurements can prove that there is a laminar film, because only velocity measurements can be used to obtain the velocity gradient, and the velocity gradient is the criterion by which the character of flow is established.

from which

$$V = \frac{2a\, r_0^m}{(2+m)(1+m)} \quad \ldots\ldots [18]$$

Combining Equation [3] and Equation [7] to eliminate the pressure gradient $-(dp)/(dx)$, the Reynolds index law becomes

$$4\,\tau_0 = A_2 \left(\frac{DV\rho}{\mu}\right)^n \frac{\mu^2}{\rho D^2} \quad \ldots\ldots [19]$$

and solving Equation [19] for V

$$V = \left(\frac{4\tau_0}{A_2\rho}\right)^{1/n} \left(\frac{\rho}{\mu}\right)^{(2-n)/n} D^{(2-n)/n} \quad \ldots\ldots [20]$$

Comparing Equation [18] with Equation [20]

$$m = (2-n)/n \quad \ldots\ldots [21]$$

Blasius had found by analysis of published experimental data that n is 1.75, which is close to the value found by Reynolds. With this value of n the value of m is $1/7$, and Equation [18] becomes

$$V = (49/60) a r_0^{1/7} \quad \ldots\ldots [18a]$$

According to Equation [18a], the ratio V to u_{max} (the velocity at the axis) is $49/60$, which is quite close to experimental values. Using the value given by Blasius for A_2, which is 0.1582, the value of a is calculated from Equation [20], and Equation [16] becomes

$$u = 8.56 \left(\frac{\tau_0}{\rho}\right)^{4/7} \left(\frac{\rho}{\mu}\right)^{1/7} y^{1/7} \quad \ldots\ldots [22]$$

Writing Equation [22] for the velocity at the axis

$$u_{max} = 8.56 \left(\frac{\tau_0}{\rho}\right)^{4/7} \left(\frac{\rho}{\mu}\right)^{1/7} r_0^{1/7} \quad \ldots\ldots [23]$$

and for the average velocity

$$V = \frac{49}{60} \cdot 8.56 \left(\frac{\tau_0}{\rho}\right)^{4/7} \left(\frac{\rho}{\mu}\right)^{1/7} r_0^{1/7} \quad \ldots\ldots [24]$$

Solving Equation [24] for τ_0

$$\tau_0 = 0.0333\, V^2 \rho \left(\frac{\mu}{r_0 V \rho}\right)^{1/4} \quad \ldots\ldots [25]$$

Dividing Equation [22] by Equation [24]

$$\frac{u}{V} = \frac{60}{49} \left(\frac{y}{r_0}\right)^{1/7} \quad \ldots\ldots [26]$$

and at the inner boundary of the surface film

$$\frac{U}{V} = \frac{60}{49} \left(\frac{b}{r_0}\right)^{1/7} \quad \ldots\ldots [27]$$

Prandtl's hypothesis, for which he offered no reasons, was that the inner boundary of the film would be located where the velocity gradient, calculated according to the turbulent-flow formula, was equal to the velocity gradient in viscous flow. Differentiating Equation [26]

$$\frac{du}{dy} = \frac{60}{343} \frac{V}{r_0} \left(\frac{r_0}{y}\right)^{6/7} \quad \ldots\ldots [28]$$

For viscous flow

$$\frac{du}{dy} = \frac{\tau_0}{\mu} \quad \ldots\ldots [29]$$

Combining Equations [28] and [29] to eliminate (du/dy), and substituting for τ_0 its value from Equation [25], there result

$$\frac{60}{343} \frac{V\mu}{r_0} \left(\frac{r_0}{b}\right)^{6/7} = 0.0333\, V^2 \rho \left(\frac{\mu}{r_0 V \rho}\right)^{1/4} \quad \ldots\ldots [30]$$

where b has been written as the value of y at the boundary.

Equation [30] simplifies to

$$\left(\frac{b}{r_0}\right)^{6/7} = \frac{60}{0.0333 \times 343} \left(\frac{\mu}{r_0 V \rho}\right)^{3/4} \quad \ldots\ldots [31]$$

and elimination of b/r_0 by combining Equations [31] and [27] yields the ratio

$$\frac{U}{V} = 1.62 \left(\frac{\mu}{r_0 V \rho}\right)^{1/8} \quad \ldots\ldots [32]$$

Prandtl did not claim high accuracy for the numerical constant 1.62; he noted that Stanton's calculations, referred to previously, which made the ratio about 0.3, corresponded to a numerical constant of about 1.2.

Equation [32] can be transformed, using the Reynolds index law with the constants given by Blasius, to

$$\frac{U}{V} = 8.85 \left(\frac{\tau_0}{\rho V^2}\right)^{1/2} \quad \ldots\ldots [33]$$

There is a slight inconsistency introduced by Prandtl's assumption, in that the velocity U calculated by Equation [13], with b taken as y and calculated by Equation [31], is much smaller than given by Equation [33]; in fact, it is only $1/7$ as great. Prandtl overcame this difficulty by changing the definition of y, making it less than the distance from the wall by 6/7 of the film thickness, that is, y is the distance from a point within the film which is (6/7) b from the wall. With this change, the value of y at the inner boundary of the film is $b/7$ rather than b, and $b/7$ should be written in place of b in Equations [27], [30], and [31], and in the accompanying text.

While Stanton's calculations indicated that Prandtl's result for the ratio U/V was too high, really good experimental data were not available until 1930, when Eagle and Ferguson (12) published the results of their very careful work on the transfer of heat from a tube to water flowing through it. To eliminate uncertainties which arise from the change of fluid properties with temperature, they made experiments at various heat-transfer rates, then extrapolated to zero heat flow.

Eagle and Ferguson also discussed the theory underlying Equation [14], which they wrote in the dimensionless form

$$\frac{\rho V C_p (t_w - t_f)}{q_0} = \left[1 + \frac{U}{V}(N_{Pr} - 1)\right] \frac{\rho V^2}{\tau_0} \quad \ldots\ldots [34]$$

They say: "It is evident that the theory by which Equation [34] is derived must be inaccurate. One obvious point in which it is incomplete is the assumption that the total thermal resistance is made up of a viscous film and a turbulent core only, each strictly obeying their quite different laws. Physically there cannot be this sharp boundary between the film and the mean motion of the turbulent core; there must be a region in which the resultant behavior of the fluid is intermediate between viscous flow and turbulence." The similarity of this reasoning to that of Prandtl quoted previously is striking, particularly in view of Stanton's remarks regarding the experimental work at the National Physical Laboratory.

Eagle and Ferguson took the view that the theory could be modified to bring in the intermediate zone, saying, "As there must be, between the film and the core, a region whose behavior is intermediate between the behavior of these, theory indicates

that $\rho\ V\ C_p\ (t_w - t_f)/q_0$, instead of being a linear function of N_{Pr} should be a continually increasing function of it with a gradient continuously decreasing to some asymptotic value, which asymptotic value would give the value of U/V for the viscous film." This reasoning was implemented by the following: "For a given value of $D\ V\ \rho/\mu$ the thermal resistance of the core and film have been seen above to be independent of N_{Pr} and directly proportional to it, respectively. Now the presence of some layer of intermediate behavior between the film and core necessitates the variation of $\rho\ V\ C_p\ (t_w - t_f)/q_0$, for a constant value of $D\ V\ \rho/\mu$, qualitatively as indicated in the accompanying figure, which shows diagrammatically the three thermal resistances of the core, film, and intermediate layer." In the figure referred to, $\rho\ V\ C_p\ (t_w - t_f)/q_0$ is plotted against N_{Pr} as the sum of three resistances; one for the core, which is constant, a second for the film, which is proportional to N_{Pr} and a third for the intermediate layer, which is proportional to the difference between N_{Pr} and a fraction of N^2_{Pr}. Eagle and Ferguson accordingly wrote

$$\rho\ V\ C_p\ (t_w - t_f)/q_0 = A_3 + B_3\ (N_{Pr} - 1) - C_3\ (N_{Pr} - 1)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [35]$$

where A_3, B_3, and C_3 are positive numbers which are functions of $DV\rho/\mu$

They found that their results could be represented very well by Equation [35], but cautioned that it might not be suitable for fluids whose N_{Pr} is very different from that of water.

The experimental results obtained by Eagle and Ferguson have been used extensively by later authors, though Equation [35] has not. In 1932 White (13) showed that Eagle and Ferguson's results were in fair agreement with a formula which he had derived on the basis of the Reynolds analogy, using reasoning very similar to that given by Eagle and Ferguson. White's formula is

$$\frac{q_0\ V}{\tau_0\ C_p\ (t_w - t_f)} = \frac{\beta}{N_{Pr}^{1/3}} \quad \ldots\ldots\ldots\ldots [36]$$

where β is a function of $D\ V\ \rho/\mu$ and N_{Pr}.

White gives a very lucid discussion of the Reynolds analogy which may be conveniently mentioned here. He points out that the definition of viscosity is

$$\tau = \mu\ (du/dz) \quad \ldots\ldots\ldots\ldots\ldots\ldots [37]$$

where

τ = resistance to flow per unit area
z = distance measured perpendicularly from the area

while the definition of thermal conductivity is

$$q = -k\ (dt/dz) \quad \ldots\ldots\ldots\ldots\ldots\ldots [38]$$

where

q = heat lost per unit area of surface perpendicular to z in unit time

In turbulent flow there is a mechanical mixing and the velocity gradients are less than Equation [37] predicts, except at the wall. For convenience, we introduce a mechanical viscosity μ_m and define it by

$$\tau = (\mu + \mu_m)\ (du/dz) \quad \ldots\ldots\ldots\ldots\ldots\ldots [39]$$

Of course μ_m depends upon the motion, and thus varies from point to point in the fluid, even if the temperature is constant, while μ is a property of the fluid itself, and remains constant if the temperature does.

In turbulent flow with transfer of heat, the temperature gradients are less than Equation [38] predicts, and, since the mixing action which accounts for the decreased velocity gradients also causes convective heat transfer, it is reasonable to introduce a mechanical thermal conductivity k_m, and define it by

$$q = -(k + k_m)\ (dt/dz) \quad \ldots\ldots\ldots\ldots\ldots\ldots [40]$$

It is plain that the quantity of momentum transferred when a certain volume of fluid changes velocity is proportional to the volume, the density of the fluid, and the change in velocity; if that same volume of fluid changes temperature, a quantity of heat is transferred which is proportional to the volume, the density of the fluid, the change in temperature, and the heat capacity of the fluid. Therefore the analog of momentum is not heat, but the ratio of heat to heat capacity; accordingly, we divide Equation [40] by C_p and obtain

$$\frac{q}{C_p} = -\left(\frac{k}{C_p} + \frac{k_m}{C_p}\right)\frac{dt}{dz} \quad \ldots\ldots\ldots\ldots [41]$$

If μ_m is very large compared to μ, and k_m very large compared to k, we may write as approximations

$$\tau = \mu_m\ (du/dz) \quad \ldots\ldots\ldots\ldots\ldots\ldots [39a]$$

and

$$q/C_p = -(k_m/C_p)\ (dt/dz) \quad \ldots\ldots\ldots\ldots [41a]$$

Reynolds analogy may then be put into a third form, that μ_m is equal to k_m/C_p, and therefore the velocity and temperature distributions should be similar. White points out that in the case of a fluid for which μ is equal to k/C_p, similarity of temperature and velocity distributions does follow from the third form of Reynolds analogy, but for other fluids this will not be the case, except under conditions such that Equations [39a] and [41a] are good approximations. He concluded that the problem was too complicated for analytical treatment.

In 1933 Colburn (14) presented a most excellent paper on the correlation of forced-convection heat-transfer data and a comparison with fluid friction, and offered the formula

$$\frac{q_0 N_{Pr}^{2/3}}{C_p V\ \rho\ (t_w - t_f)} = \frac{\tau_0}{\rho\ V^2} \quad \ldots\ldots\ldots\ldots [42]$$

Equation [42] is very similar to Equation [36], and would be the same as Equation [36] if β were $1/N_{Pr}^{1/6}$. Colburn also discussed the Reynolds analogy and the development of Equation [14], and pointed out that Eagle and Ferguson had considered a buffer zone between the laminar film and the turbulent core. Colburn concluded that heat-transfer theory for turbulent-flow conditions had not further advanced because of the lack of simple mathematical relationships for turbulent flow itself.

The problem was taken up at about the same time by von Kármán (15) who published one paper in 1934, and two others in 1939. Von Kármán explained the Reynolds analogy very much as White had, and adopted the Eagle and Ferguson criticism of the theory that $\rho\ V\ C_p\ (t_w - t_f)/q_0$ is a linear function of N_{Pr}. He also adopted[7] their scheme of dividing the thermal

[7] In a discussion of the 1941 paper by Boelter, Martinelli, and Jonassen (16) von Kármán wrote: "The new expression 'buffer layer' for the transition layer between perfectly laminar and perfectly turbulent regions, introduced by the writer in order to explain the discrepancies between experiments and the theories of Prandtl and G. I. Taylor, is quite appropriate." This indicates that von Kármán independently conceived the existence of a transition zone between the viscous film and the turbulent core, and believed the conception to be original with him. Since he used the experimental data of Eagle and Ferguson, and cites their work in his bibliography, this is at first glance astonishing. However, in a footnote in the 1934 paper, von Kármán states that he obtained the experimental results he used from a then-current review by H. Lorentz, so it may be that he never read the Eagle and Ferguson discussion. Later authors have erroneously credited von Kármán with the basic concept of dividing the tube cross section for the purpose of analysis into three

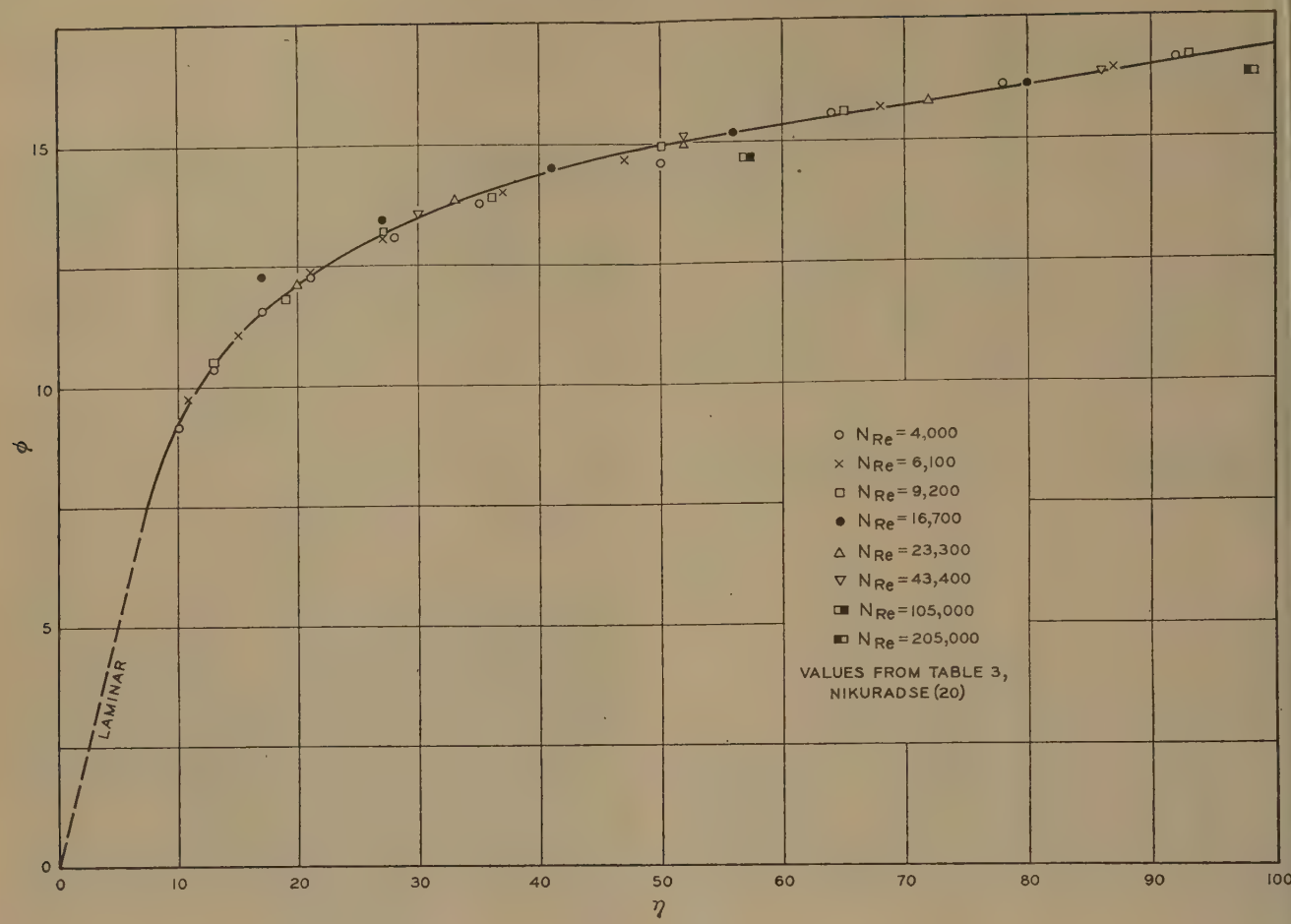


FIG. 1 GENERALIZED VELOCITY DISTRIBUTION [AFTER NIKURADSE (20)]

resistance into three parts: one for the turbulent core, which is independent of N_{Pr}; one for the laminar film, which is a linear function of N_{Pr}; and one for the transition region, which increases with N_{Pr}, but more slowly than the first power.

The novel contribution made by von Kármán was this: He proposed to use a velocity distribution published by Nikuradse (20) to obtain the thickness of the film and of the transition layer, and the rate of increase of thermal resistance in the transition layer with N_{Pr}; thus he would predict heat-transfer rates from velocity measurements, as Stanton had proposed but had been unable to accomplish. Eagle and Ferguson had obtained the constants in their formula from heat-transfer measurements. Unfortunately, the numerical values published by Nikuradse and used by von Kármán were not experimental values, but values adjusted to bring them into agreement with the laminar-film hypothesis.

Nikuradse described experimental work which he had done on flow of water through drawn brass tubes, including velocity distributions measured in sixteen experiments in which N_{Re} (the Reynolds number, which is $D\ V\ \rho/\mu$) varied from 4000 to 3,240,000. These measurements he proposed to use to test various theories of turbulent flow. One of these, which Nikuradse ascribed to Prandtl, was that the velocity at a point close to the wall should depend only upon the resistance to flow per unit area of wall, the viscosity and the density of the fluid, and of course the distance from the wall, and should be independent of the distance to the opposite wall and of the average or maximum velocity; that is, u should be a function of τ_0, μ, ρ, and y, but independent of the size or shape of the conduit, of V, and of u_{max}.

Since $(\tau_0/\rho)^{1/2}$ has the dimensions of velocity, while $\mu/(\tau_0\ \rho)^{1/2}$ has the dimension of length, Prandtl's theory was that the ratio of u to $(\tau_0/\rho)^{1/2}$ should be a function of the ratio of y to $\mu/(\tau_0\ \rho)^{1/2}$. The first ratio Nikuradse called ϕ and the second he called η.

Nikuradse published a table (his Table 3) which purportedly gave the values of ϕ and η calculated from his measurements, and also two figures, in which ϕ was plotted against η (his Fig. 23) and against $\log_{10}\eta$ (his Fig. 24). The accompanying Fig. 1 is patterned after a portion of Nikuradse's Fig. 23, including the dashed line marked "laminar," which is plotted for ϕ equal to η; this is approximately correct near the wall for laminar flow.

Nikuradse's Fig. 24 included three lines. One was a curved line, said to be calculated on the one-seventh-power assumption (Equation [33], very nearly); this line need not concern us. Then there were two straight lines. One of these is given by

$$\phi = 5.75 \log_{10} \eta + 5.5 \qquad [43]$$

Nikuradse said that this line represents the points close to the axis of the tube. The other straight line is given by

$$\phi = 5.52 \log_{10} \eta + 5.84 \qquad [43a]$$

Nikuradse said that the second line represents the points near the wall of the tube.

Fig. 2 is patterned after a portion of Nikuradse's Fig. 24, but drawn as suggested by von Kármán, and as it has appeared with minor variations in various heat-transfer publications. It in-

regions; a viscous film, a turbulent core, and a buffer zone (16, 17, 18, 19).

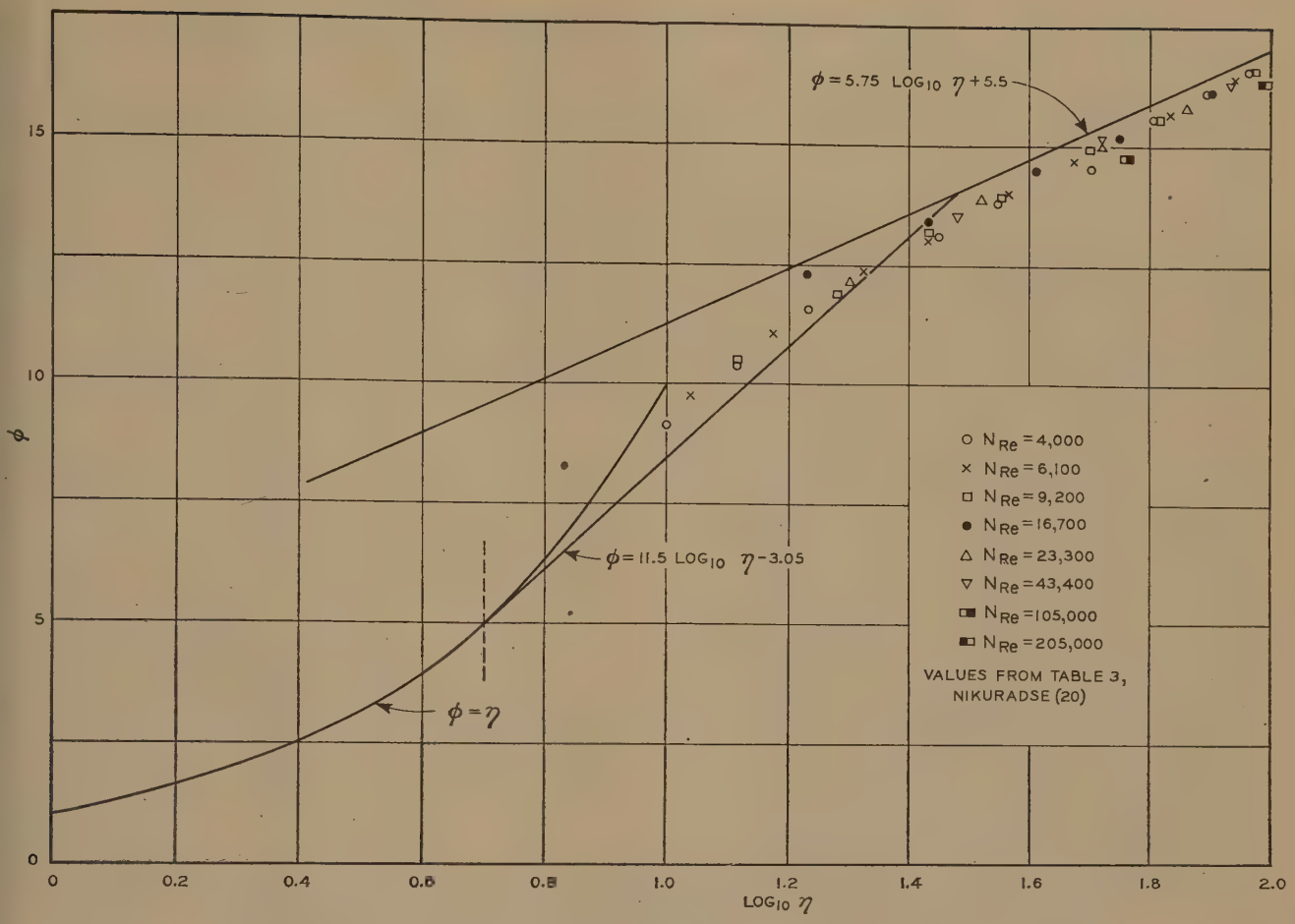


FIG. 2 GENERALIZED VELOCITY DISTRIBUTION [AFTER NIKURADSE (20) PLUS VON KÁRMÁN (15)]

cludes the straight line of Equation [43], but not the straight line of Equation {43*a*}, because von Kármán used only the former; it includes also the line for ϕ equal to η which is not in Nikuradse's Fig. 24, but is in his Fig. 23; and it includes a straight line given by

$$\phi = 11.5 \log_{10} \eta - 3.05 \quad \ldots\ldots\ldots\ldots\ldots [44]$$

The line given by Equation [44] does not appear in Nikuradse's publication, but was drawn by von Kármán.

The theory of von Kármán is that Equation [44] is based upon Nikuradse's measurements, and represents the velocity distribution in the transition region between laminar and fully turbulent flow, and that the intersection of the line of Equation [44] with the line which corresponds to ϕ equal to η marks the inner boundary of the laminar film, while the intersection of the line of Equation [44] with the line of Equation [43] marks the inner boundary of the transition region.

The author had occasion to examine Nikuradse's paper carefully, and to recalculate some of the numerical values given in his Table 3 from the velocity distributions themselves, which are given in his Table 2. It became apparent that there was a serious discrepancy between Table 2 and Table 3, in that each value of η in Table 3 was greater by 7 than the corresponding value calculated from the data in Table 2. When the values of ϕ were plotted against the values of $\log_{10} \eta$ calculated from the data in Table 2, the result was Fig. 3 of this paper. It will be noted that all the points in Fig. 3 lie close to a straight line, and there is no suggestion of any trend toward the line for ϕ equal to η.

Since the adjusted values of Nikuradse were being used by writers on heat transfer who followed von Kármán, it seemed desirable to call this discrepancy to their attention: It was obvious that if Table 2 is consistent with Nikuradse's measurements and Table 3 is not, there is no experimental basis for the velocity distribution von Kármán used in his heat-transfer calculations. It was suggested, however, that Table 2 contains the raw values of Nikuradse's observations, which had to be corrected for the interference between the wall of the pipe and the Pitot tube used for the measurements, while Table 3 contains the corrected values.

Nothing could be found in Nikuradse's publication to support this suggestion; on the contrary, Nikuradse stated that, to avoid such interference, the Pitot tube was not placed within the pipe at all, but instead the measurements were made in the free jet just downstream from the end of the pipe. There was a discussion of this matter in December, 1946, in which Professor Rouse, Professor Boelter, and Professor Martinelli participated, on which occasion Professor Rouse suggested that it would be possible to inquire of Professor Prandtl concerning the discrepancy; since Nikuradse had worked under Prandtl's direction, this was a particularly happy suggestion. Accordingly, an inquiry was prepared by the author, and sent to Prandtl by Professor Rouse through the United States Government intelligence agency working in Germany, known as FIAT. The reply to this inquiry explained the discrepancy.

It appeared that of recent years Nikuradse had been an instructor at the Technical University in Breslau; as the Soviet Army had occupied Breslau, Nikuradse was visiting relatives in Göttingen, the city in which the Kaiser Wilhelm Institute for

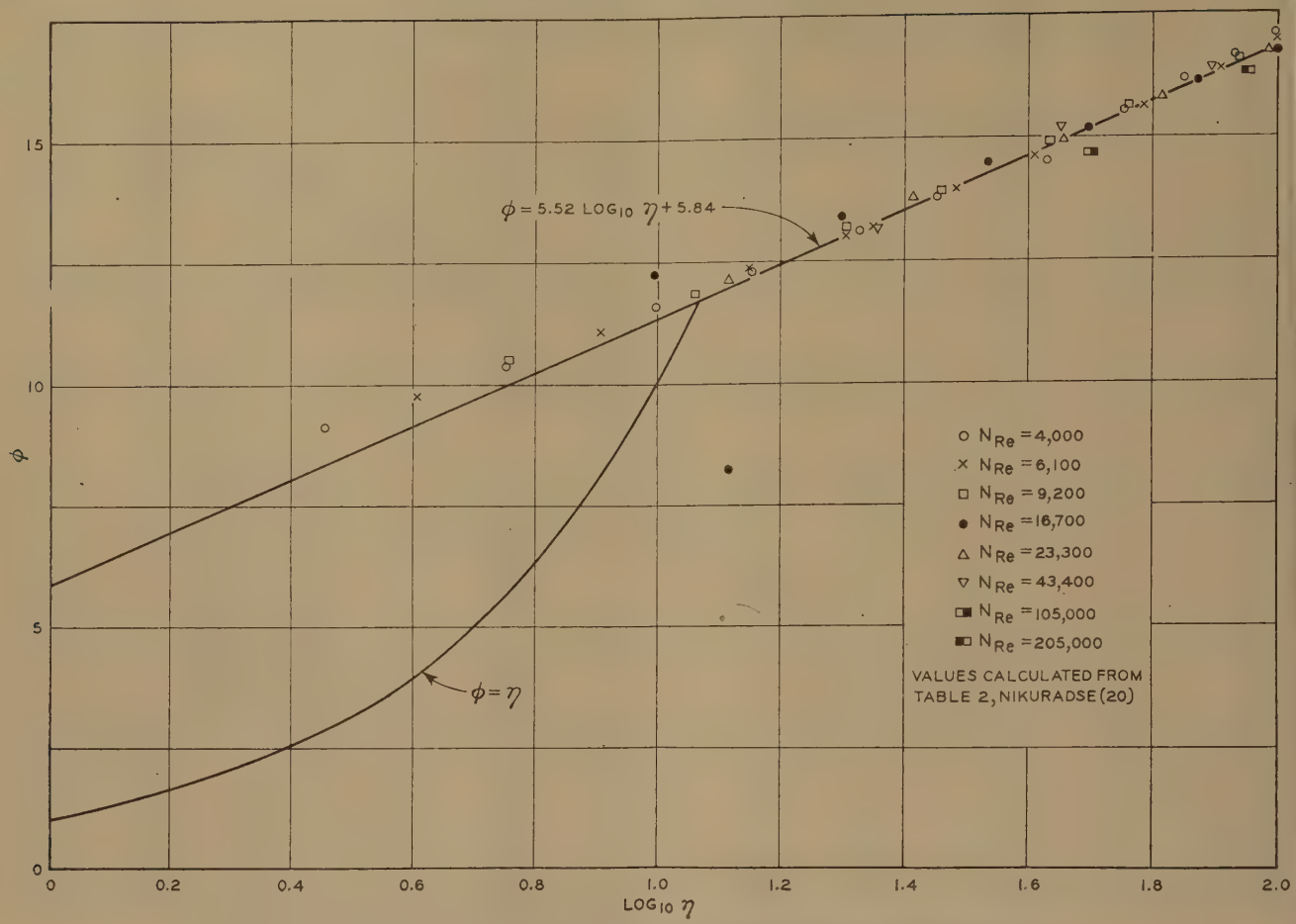


FIG. 3 GENERALIZED VELOCITY DISTRIBUTION, FROM NIKURADSE'S EXPERIMENTAL DATA

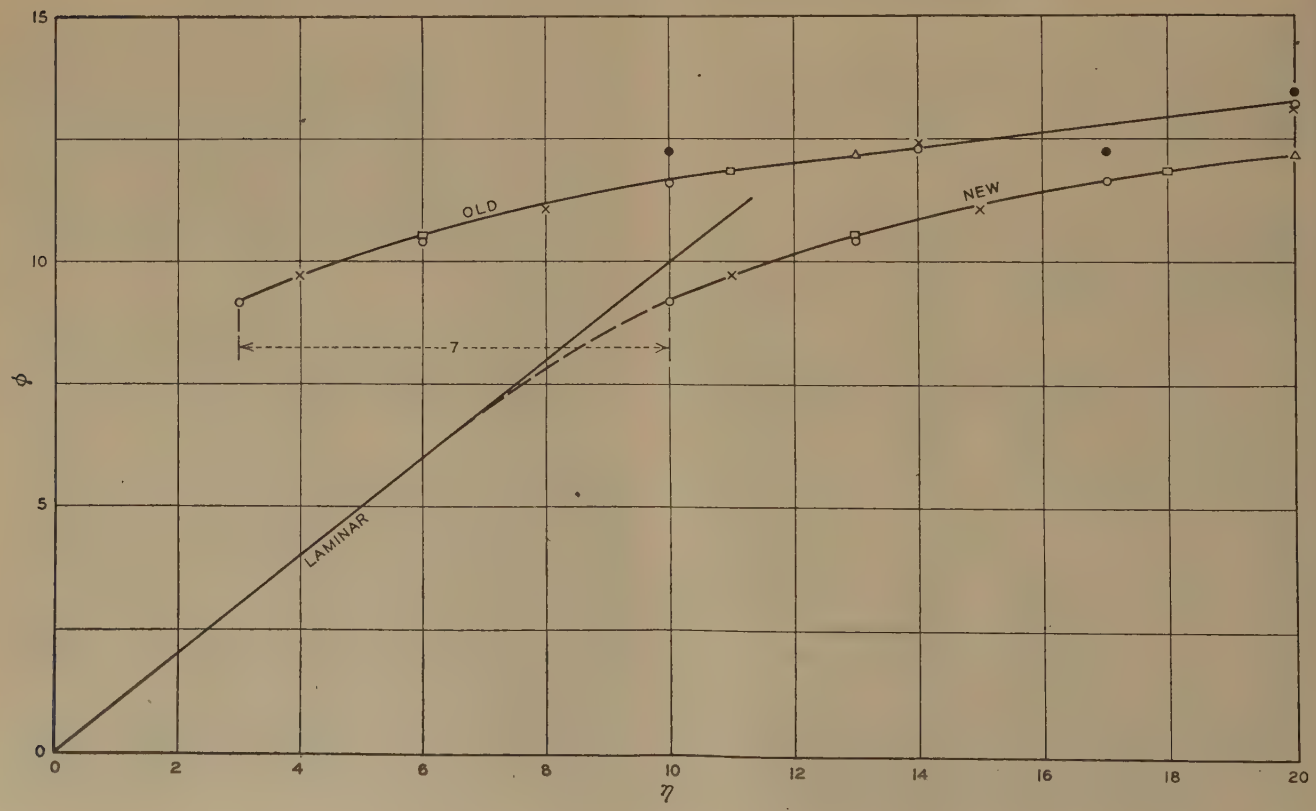


FIG. 4 PRANDTL'S EXPLANATION OF THE NIKURADSE SHIFT

Flow Research is located, at the time the inquiry arrived. This is the institute in which Nikuradse made his velocity measurements; Prandtl was the director of this institute at the time, and had returned to it. Therefore it was possible for Prandtl and one of his colleagues to interview Nikuradse, and find out what had happened.

According to Prandtl, after Nikuradse had prepared several of his tables, he plotted ϕ against η using his measured values. The result is shown in Fig. 4 herewith, the points lying along the line marked "old" being the measured values. Then he plotted the straight line for ϕ equal to η which is marked "laminar." As can be seen, the line marked laminar would, if prolonged, intersect the line marked old, which means that some of the velocities measured were greater than they would have been if the flow were laminar rather than turbulent. This was not reasonable, so Nikuradse decided that his measurements were not so exact as he had first assumed.

If there is a laminar film, the curve for the turbulent portion of the flow should become tangent to the straight line for laminar flow at the film boundary. To get a curve for the turbulent portion of the flow which is tangent to the straight line, Nikuradse added 7 units to each value of η which shifted each point 7 units to the right. The line drawn through the shifted points is marked "new" in the figure.

The line marked new in Fig. 4 is tangent to the straight line at about η equal to 7. However, if Nikuradse had chosen to use a different French curve, he would have achieved tangency at some other value of η, and the number of units in the shift would have been different. He might have made his shifted curve tangent to the straight line at η equal to zero, in which case his published values would have "established" that there is no laminar film!

It must be emphasized that there was absolutely no basis for the magnitude of the shift in the values of η. This is not a case such as frequently occurs when there is an error in observation, and the magnitude of the error is estimated on the basis of calibration experiments, and then an adjustment made to correct the observed data. The values of η depend upon the distance from the wall to the point at which the velocity is to be measured, and the placement of the Pitot tube had been carefully worked out by means of preliminary experiments. The shift of 7 units was simply a distortion of the data to make it seem to agree with an hypothesis.

However, Nikuradse did not claim to have proved the existence of the laminar film by his velocity measurements. Furthermore, his measured values of η ranged up to 55,000, and by far the greater part of them were large enough so that the addition of 7 units was an insignificant distortion. Nikuradse made no use of the altered η values in his further calculations. If Nikuradse's published values had not been used by later investigators, his shift would have had little significance.

The use of Nikuradse's published values in heat-transfer calculations gave them an importance which they might not otherwise have had. As Stanton pointed out, if heat-transfer rates calculated by the Reynolds theory using an experimental velocity distribution should be found to agree with experimental heat-transfer rates, the Reynolds' theory would gain considerable support. It is of course possible to calculate a velocity distribution which would make heat-transfer measurements agree with the Reynolds' theory, but that would be no proof that the theory is correct.

Boelter, Martinelli, and Jonassen (16) calculated such a velocity distribution, that is, one which would make heat-transfer measurements agree with the Reynolds theory. They followed von Kármán in accepting Nikuradse's equation (given as Equation [43]) for the velocity distribution in the core where the flow is assumed to be fully turbulent; then they assumed that the velocity distribution in the intermediate zone could be approximated by a straight line on the plot of ϕ against $\log_{10} \eta$, and that μ_m is equal to k_m/C_p, and determined what straight line would give the best agreement with heat-transfer data on the basis of these assumptions.

The heat-transfer data used by Boelter, Martinelli, and Jonassen were obtained from experiments in which very little of the heat transferred from the wall to the fluid reached the center of the pipe, so that most of the resistance to heat flow was attributed to the laminar film and the intermediate zone, and only a minor part to the turbulent core. For this reason the assumption that in the core the velocity distribution is given by Equation [43], is relatively unimportant, and a quite different assumption would not have changed their finding materially.

This finding was that von Kármán's line, given by Equation [44], gave the best agreement with the heat-transfer data they used. This is a remarkable finding, as will be apparent on looking again at Fig. 2. The figure shows that the points lie above von Kármán's line, and the ordinary investigator would probably have drawn the line so as to intersect the curve for ϕ equal to η at a value of ϕ greater than 8, and such a line would not have been confirmed by the heat-transfer measurements.

Martinelli (18) calls attention to the coincidence between the finding of Boelter, Martinelli, and Jonassen based upon heat-transfer data and von Kármán's representation of Nikuradse's pseudo-observations, but points out that the Reynolds theory has been modified by later investigators who assume that μ_m is proportional to k_m/C_p, with the proportionality constant lying between 1 and 2, and that if the proportionality constant is taken as greater than 1, the thickness of the laminar film and of the intermediate zone calculated from heat-transfer data would no longer correspond to those which von Kármán obtained from Nikuradse's "measurements."

It has been suggested that the existence of the laminar film, its thickness, and the velocity distribution just outside have been established by measurements reported by Reichardt (21). The publication did not describe completely how the measurements were made, but what was published suggests caution in accepting the results. In the first place, the flow was not in a round tube, but at the end of a rectangular channel 25 cm high and 100 cm wide. It is still to be proved that the velocity near a curved wall follows the same law as the velocity near a flat wall, and in the last analysis, measurements near a curved wall are needed to test the hypothesis. Then the velocities were measured with hot wires, which had been calibrated in a rectangular channel 3 cm high and 30 cm wide through which air was passing in laminar flow, the distances from the wall in the calibration tests being chosen so as to get the same value of resistance to flow as would exist in the turbulent stream of the large channel. The assumption implied in this method of calibration is still to be proved. Finally, the measured pressure drop was higher than that calculated on the basis of measurements of other authors, and therefore the value of τ_0 used in the calculations was not the measured value, but one estimated on the basis of Nikuradse's report.

It seems that the presence of the viscous film on the wall of a pipe through which fluid is flowing turbulently is still to be demonstrated by actual velocity measurements. The difficulty of making such measurements is not to be minimized.

The basic hypothesis is not without difficulties on the theoretical side. For one, the meaning of "thickness of the laminar film" requires further discussion. It is not reasonable, as Prandtl pointed out, to imagine a sharp boundary between laminar and turbulent regions. Even if we consider that the eddies gradually become smaller as the wall of the pipe is approached, it cannot be

that there is a definite boundary which the eddies can never penetrate. Instead, the thickness of the laminar film must be a statistical thing, a kind of average of values which differ from place to place over the surface of the pipe, and from instant to instant at each place, these values ranging from zero upward.

If the thickness of the laminar film varies with time and place, the effective value of film thickness for momentum transfer need not be the same as the effective value for heat transfer. If the thermal conductivity were zero, for example, no heat would be transferred at all through a finite laminar film. On the other hand, if the film thickness were sometimes zero, heat could be transferred even if k were zero. Therefore it is suggested that the effective laminar-layer thickness for heat transfer may be greater or less than it is for momentum transfer, depending upon N_{Pr}; when N_{Pr} is high, the ratio of heat-transfer thickness to momentum-transfer thickness is low, and vice versa. This is particularly important in the case of fluids of high N_{Pr}, since the thermal resistance of the film is most important with poorly conducting fluids.

The idea that the effective film thickness for heat transfer decreases as N_{Pr} increases comes primarily from a consideration of empirical formulas for heat transfer, such as Colburn's, given as Equation [42]. According to such formulas the entire thermal resistance is proportional to the Prandtl number raised to a power less than 1. Such empirical formulas represent test data on poorly conducting fluids well enough to be used for design purposes, and may be said to have a good basis in observations; the laminar-film hypothesis should be capable of accounting for such observations.

As pointed out previously, the thermal resistance of the laminar film is directly proportional to the Prandtl number. Suppose we consider heat transfer to a large number of fluids having the same heat capacity and viscosity, but varying in thermal conductivity. According to the empirical formulas, the total thermal resistance will increase with the resistivity of the fluid, but less rapidly, i.e., with the $^2/_3$ power of the fluid resistivity, according to Colburn. The resistance of unit thickness of the laminar film itself increases directly with the fluid resistivity, so that with constant laminar film thickness we soon reach the point that the resistance of the film alone is greater than the total resistance, no matter how thin the laminar film may be assumed to be.

At present only the sketchiest ideas are available concerning a mechanism which would permit the effective thickness of the laminar film for heat transfer to be less than the effective thickness for momentum transfer. One might imagine that where an eddy partially penetrated the laminar film, both heat and momentum would be transferred across the remaining film thickness, while heat alone could be transferred through the sides of the eddy between the body of the eddy and the adjacent portions of the laminar film.

In case the fluid has high thermal conductivity the convective resistance is relatively more important. Then the error due to using Equations [39a] and [41a] becomes appreciable, and even the scheme of dividing the thermal resistance into three parts is too crude. It is necessary to deal with k_m as a continuously variable function of the distance from the pipe axis, one which has a value of zero at the laminar layer inner boundary, and which may be small compared to k even where the turbulence is greatest. Some progress has already been made in this direction.

Martinelli (18) has discussed heat transfer in molten metals where electron convection as well as molecule convection contribute to the thermal conductivity, making it comparable to the mechanical conductivity even at high turbulence. His paper well shows the imperfections resulting from the basic weaknesses of the logarithmic expressions—Equations [43] and [44]—and mentions the possibility of improvement by using an expression fo the velocity distribution developed by the present author.

Another difficulty with the basic hypothesis appears when w consider heat transfer between a tube and a gas flowing throug it at low pressure and high velocity. Let us take as an example tube 1 cm in diameter through which is flowing at a velocity o 27,000 cm per sec a gas whose viscosity is 1.5×10^{-4} poise an whose density is 8.8×10^{-5} g per cu cm, the pressure being 50,00 dynes per sq cm and the temperature 200 deg K.

The Reynolds number of the flow, $D\ V\ \rho/\mu$, is $(1 \times 2.7 \times 10^4 \times 8.8 \times 10^{-5})/(1.5 \times 10^{-4})$, or 1.6×10^4, so that the flow i turbulent. At a Reynolds number of 1.6×10^4 the ratio of V t $(\tau_0/\rho)^{1/2}$ is 17, whence $(\tau_0/\rho)^{1/2}$ is 27,000/17, or 1600. Multi plying by ρ, which is 8.8×10^{-5}, $(\tau_0\ \rho)^{1/2}$ is found to be 0.14 Then the value of $\mu/(\tau_0\ \rho)^{1/2}$ is $1.5 \times 10^{-4}/0.14$, which is 1.1×10^{-3}, and η is $y/(1.1 \times 10^{-3})$.

Accepting for this example the finding of Boelter, Martinelli and Jonassen (16) that the laminar-film boundary is at η equal t 5, we have that b, the value of y at the laminar-film boundary, i 5.5×10^{-3} cm, or the thickness of the laminar film is $5.5 \times 10^-$ cm.

According to kinetic theory, the mean free path of the mole cules of a gas, denoted by λ, is given approximately by

$$\lambda = 1.5\ \mu/(p\ \rho)^{1/2} \quad \text{[45]}$$

from which in this example λ is $(1.5 \times 1.5 \times 10^{-4})/(5 \times 10^ \times 8.8 \times 10^{-5})^{1/2}$, or 1.1×10^{-4} cm.

It thus appears that the thickness of the laminar film is only about 50 times the mean free path of the gas molecules for the conditions of this example. Whether the conduction of heat through a gas for a distance approaching in magnitude the length of the mean free path of the molecules may be calculated using the thermal conductivity measured in the usual way is questionable.

BIBLIOGRAPHY

1 "Traité de la Chaleur," by Eugène Péclet, V. Masson, Paris, third edition, vol. 1, 1860, pp. 385–395; an English translation of these paragraphs is in "Condensation of Steam in Covered and Bare Pipes," by C. P. Paulding, D. Van Nostrand Company, Inc., New York, N. Y., 1904.

2 Article by O. Reynolds, Proceedings of Literary and Philosophical Society of Manchester, England, vol. 13, 1874, p. 7; reprinted in (10).

3 "On the Passage of Heat Between Metal Surfaces and Liquids in Contact With Them," by T. E. Stanton, Philosophical Transactions, Royal Society of London, series A, vol. 190, 1897, pp. 67–88.

4 "Steam Engine," by J. Perry, The Macmillan Company, London, England, 1904, pp. 595–596.

5 "Eine Beziehung zwischen Wärmeaustausch und Strömungswiderstand der Flüssigkeiten," by L. Prandtl, *Physikalische Zeitschrift*, vol. 11, 1910, pp. 1072–1078.

6 "Conditions at the Surface of a Hot Body Exposed to the Wind," by G. I. Taylor, Great Britain Advisory Committee for Aeronautics, Reports and Memoranda No. 272, 1916.

7 "Heat Transmission Over Surfaces," by T. E. Stanton, Great Britain Advisory Committee for Aeronautics, Reports and Memoranda No. 243, 1916.

8 "On the Conditions at the Boundary of a Fluid in Turbulent Motion," by T. E. Stanton, Dorothy Marshall, and Mrs. C. N. Bryant, Proceedings of the Royal Society of London, England, series A, vol. 97, 1920, pp. 413–434.

9 "Bemerkung über den Wärmeübergang in Rohr," by L. Prandtl, *Physikalische Zeitschrift*, vol. 29, 1928, pp. 487–489.

10 "Das Aehnlichkeitsgesetz bei Reibungsvorgängen in Flüssigkeiten," by H. Blasius, *Mitteilungen über Forschungs Arbeiten a.d. Ingenieur Wesens*, Nr. 131, Berlin, Germany, 1931, pp. 1–39.

11 "Über laminare und turbulente Rebung," by T. von Kármán, *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 1, 1921, pp. 233–252.

12 "The Coefficients of Heat Transfer From Tube to Water," by Albert Eagle and R. M. Ferguson, Proceedings of The Institution of

Mechanical Engineers, England, vol. 2, 1930, pp. 985–1030; also Proceedings of the Royal Society of London, England, series A, vol. 127, 1930, pp. 540–566.

13 "Fluid Friction and Its Relation to Heat Transfer," by C. M. White, Trans. Institution of Chemical Engineers, London, England, vol. 10, 1932, pp. 66–80.

14 "A Method of Correlating Forced Convection Heat Transfer, Data and a Comparison With Fluid Friction," by Allan P. Colburn, Trans. American Institute of Chemical Engineers, vol. 29, 1933, pp. 174–209.

15 "Aspects of Turbulence Problems," by T. von Kármán, Proceedings of the Fourth International Congress for Applied Mechanics, Cambridge, England, 1934.

"Analogy Between Fluid Friction and Heat Transfer," by T. von Kármán, Trans. ASME, vol. 61, 1939, pp. 705–710; also *Engineering*, England, vol. 148, 1939, pp. 210–213.

16 "Remarks on the Analogy Between Heat Transfer and Momentum Transfer," by L. M. K. Boelter, R. C. Martinelli, and Finn Jonassen, Trans. ASME, vol. 63, 1941, pp. 447–454.

17 "Temperature Gradients in Turbulent Gas Streams," by W. H. Corcoran, B. Roudebush, and B. H. Sage, Trans. AIChE (Chemical Engineering Progress), vol. 43, 1947, pp. 135–142.

18 "Heat Transfer to Molten Metals," by R. C. Martinelli, Trans. ASME, vol. 69, 1947, pp. 947–956.

19 "Heat Transmission," by W. H. McAdams, McGraw-Hill Book Company, Inc., New York, N. Y., 1942.

20 "Gesetzmässigkeiten der turbulenten Strömung in glatten Rohren," by J. Nikuradse, Forschungs-Heft 356, Berlin, Germany, 1932.

21 "Die Wärmeübertragung in turbulenten Reibungsschichten," by H. Reichardt, *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 20, 1940, pp. 297–328; translated in NACA Technical Memo No. 1047, 1943, but it is believed that the meaning of certain parts of the original may differ from the translation.

A Study of Three Tube Arrangements in Unbaffled Tubular Heat Exchangers

By O. P. BERGELIN,[1] E. S. DAVIS,[2] AND H. L. HULL[3]

Pressure-drop and heat-transfer tests, using a medium-viscosity oil in viscous flow across vertical tubes in three unbaffled exchangers, are reported in the course of a research program on tubular exchangers. The three once-flow-through units have, respectively, equilateral triangular, in-line square, and staggered square tube arrangements, but, in all other respects, are as nearly identical as possible. Heat-transfer data for three bulk-oil temperatures and a constant, but lower, tube-wall temperature, and both isothermal and nonisothermal pressure-drop data at three temperature levels are reported. The two staggered-tube arrangements provide greater heat transfer at a given pumping power loss than the in-line arrangement in the region of viscous flow, but this superiority decreases at higher velocities and may disappear in turbulent flow as previous data have indicated. A correction is presented for the effect of viscosity gradient upon friction during heat transfer but this correction does not allow for free convection or nonuniform flow.

Nomenclature

The following nomenclature is used in the paper:

A = outer surface area of tubes, sq ft
b = longitudinal pitch, distance between center lines of adjoining transverse rows, ft
c = heat capacity, Btu/(lb)(deg F)
D_e = equivalent diameter = $\frac{4 \times \text{free volume}}{\text{exposed area of tubes}}$, ft
E = pumping power loss per unit heat-transfer area, (ft)(lb)/(hr)(sq ft)
f = friction factor defined by $\frac{f}{2} = \frac{g_c \Delta p \rho D_e}{4 G_m^2 L}$, dimensionless
G_m = mass velocity through minimum cross section, lb/(hr)(sq ft)
g_c = conversion factor, 4.18×10^8 (mass lb)(ft)/(force lb)(hr)^2
h = surface coefficient of heat transfer based on outside surface area of tubes, Btu/(hr)(sq ft)(deg F)
j = heat-transfer factor, $j = \left(\frac{h}{cG_m}\right)(N_{Pr})^{2/3}\left(\frac{\mu_s}{\mu}\right)^{0.14}$, dimensionless
k = thermal conductivity, Btu/(hr)(sq ft)(deg F/ft)
p = pressure, psf
L = length of flow, $L = Nb$, ft
N = number of rows of tubes in direction of flow
N_{Pr} = Prandtl number, $\frac{c\mu}{k}$, dimensionless
N_{Re} = Reynolds number, $\frac{D_e G_m}{\mu}$, dimensionless
n = a constant
W = mass rate of flow, lb per hr
μ = absolute viscosity of fluid at average bulk temperature, lb/(hr)(ft)
μ_s = absolute viscosity at tube-surface temperature, lb/(hr)(ft)
ρ = density, lb per cu ft

Introduction

Although the arrangement of tubes with respect to the shell-side fluid, i.e., whether "staggered" or "in-line," is important for heat transfer and fluid friction in heat exchangers, very little work has been done to compare the different arrangements. The chief available work on this subject is that of Pierson, et al (1),[4] which was carried out with air as the shell-side fluid, and covers only the region of turbulent flow. In the course of an intensive study of heat exchangers being carried out in these laboratories, it therefore seemed desirable to put the early stress on a comparison of different arrangements for the region of viscous flow as exists with medium and high-viscosity oils, especially where small tubes are used.

In a previous paper (2) data were reported for cooling a medium-viscosity oil in a once-flow-through unit, composed of $^3/_8$-in. tubes on an equilateral triangular tube arrangement with a pitch-to-diameter ratio of 1.25. The present paper compares these results with data on two additional units containing the same size tubes with the same pitch, but in one the tubes have an in-line square arrangement, and in the other they have a staggered square arrangement. All three arrangements are shown in the tube-sheet layouts given in Fig. 1. Further tests will deal with different tube sizes and different pitches.

The previous work included data under isothermal and nonisothermal conditions at one temperature level; the present study includes isothermal and nonisothermal (oil-cooling) data at three bulk-temperature levels and with essentially a constant tube-wall temperature. These results are part of a systematic study to develop a procedure to allow for nonconstant physical properties. The complete program calls for tests at other bulk temperatures and other surface temperatures, including cases of oil heating.

Apparatus, Procedure, and Results

The general apparatus was briefly described in the previous paper (2), which also included the original data on the equilateral-triangular tube arrangement at one temperature level. Further details on the apparatus and methods of calculation, and the original data on the two square arrangements, as well as data at other temperature levels on the triangular arrangement, are given in a forthcoming bulletin (3).

[1] Associate Professor, Department of Chemical Engineering, University of Delaware, Newark, Del.
[2] Research Fellow in Chemical Engineering, University of Delaware. Present address, E. I. du Pont de Nemours and Company, Wilmington, Del.
[3] Research Fellow in Chemical Engineering, University of Delaware.

Contributed by the Heat Transfer Division and presented at the Annual Meeting, New York, N. Y., November 28–December 3, 1948, of The American Society of Mechanical Engineers.

 Paper No. 48—A-34.

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

Pertinent dimensions of the three units are given in Table 1. It will be seen from this table as well as from the tube layouts

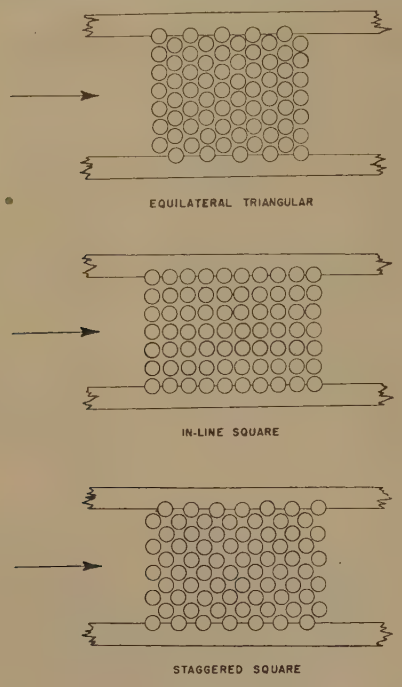


FIG. 1 TUBE ARRANGEMENTS

shown in Fig. 1 that, while the same number of tubes, i.e., 70, is used in each unit, the effective number of tubes as regards to exposed surface is 65, 60, and 63, for the triangular, in-line, and staggered square arrangements, respectively, because of the means used to minimize the wall effect. During tests the unit is mounted with the tubes vertical as shown in Fig. 2 which is a view of the in-line square unit. Tubes $^3/_8$ in. diam and 6 in. long, spaced on a 1.25 pitch-to-diameter ratio, are used in all three units.

TABLE 1 EXCHANGER DIMENSIONS

DIMENSIONS COMMON TO ALL EXCHANGERS

Outside tube diameter, in.	0.375
Number of 18 BWG copper tubes	70
Ratio of pitch to outside tube diameter	1.25
Exposed length of tube, in.	6.0
Diameter of cores inside tubes, in.	0.1875

IMPORTANT DIMENSIONAL DIFFERENCES

	Triangular	In-line, square	Staggered, square
Number of tube rows	10	10	14
Equivalent exposed tubes	65	60	63
Shell-side minimum flow area, sq in.	3.66	3.37	5.07
Shell-side equivalent diameter, in.	0.270	0.371	0.371
Shell-side heat-transfer area, sq in.	460	424	446
Shell-side friction area, sq in.	496	447	485
Length of tube bank, 12 *Nb*, in.	4.06	4.69	4.65

The oil flowing outside the tubes is a highly refined turbine oil with a viscosity of 525 SUV at 100 F. Full data on the physical properties of the oil are given in the previous publication (2), as well as in the forthcoming bulletin (3). During cooling tests water is used inside the tubes at a high velocity, i.e., cores are used to obtain a water velocity of around 7 fps, with a resulting water-film coefficient of heat transfer of approximately 1500. A "standard" test run is made with each unit upon installation, and this is repeated as the final test for each model. In every case, the initial results have been duplicated by the check run, showing that no appreciable fouling of the unit occurred during the period of testing.

Only data obtained with a heat balance of better than ±5 per cent were considered acceptable, except in cases of oil velocities less than 0.2 fps, where the temperature change of the water stream was too small to give reliable heat balances. For these cases the heat load was calculated from the oil stream, and all conditions were maintained as constant as possible so there is reason to believe that the data are of an accuracy comparable to those at higher oil velocities.

DISCUSSION OF RESULTS

Fluid Friction. The isothermal data at 150 F are shown in Fig. 3 as a simple plot of pressure drop versus the linear velocity of the oil through the minimum cross-sectional area for flow. The data for the in-line square arrangement fall along the lower line, while the upper line represents both of the staggered arrangements. Although the data from the staggered units can be represented by a single line, this agreement is believed to be only a coincidence because the triangular arrangement has a depth of 10 rows of tubes compared to 14 for the staggered square, and the amount of surface exposed to the stream is different for the two arrangements. The curves in Fig. 3 have a slope of unity at low velocities, indicating viscous flow, but above a velocity of approximately 1 fps, the slopes begin to increase, showing an increasing amount of turbulence in the tube bank. This pressure drop versus velocity plot is of value in showing the consistency and trend of the data but does not represent a satisfactory correlation because separate lines are obtained for each tube arrangement and also for each temperature level.

In Fig. 4 the isothermal data at three temperature levels are shown as friction factor versus the Reynolds number, according to the method of Chilton and Genereaux (4). This method of presentation evidently correlates the effect of the change in fluid properties with temperature but does not take care of variation in tube arrangement. The lines are straight and parallel in the viscous-flow region, but the variations in the rate of curvature in

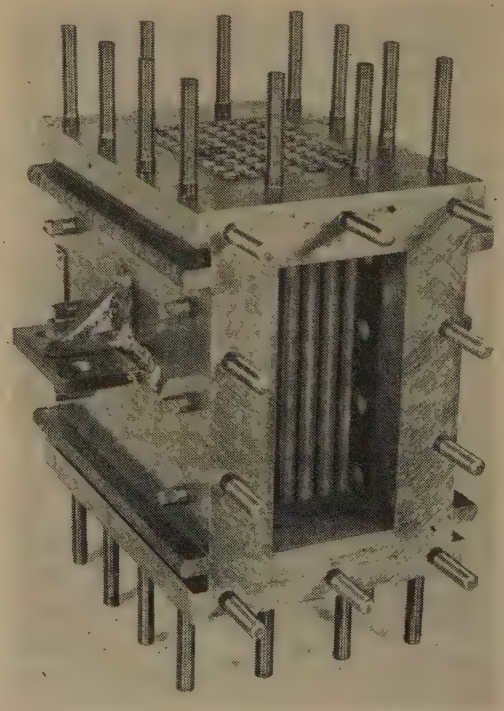


FIG. 2 MODEL EXCHANGER WITH IN-LINE SQUARE ARRANGEMENT OF TUBES

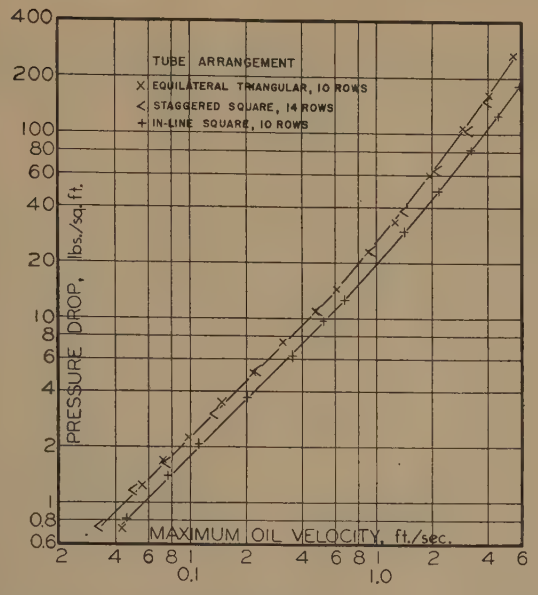


FIG. 3 ISOTHERMAL PRESSURE DROP AT 150 F FOR THREE TUBE ARRANGEMENTS

(Flow normal to 3/8-in-OD vertical tubes on 15/32-in. pitch. Oil-temperature range: 147–152 F. Gulfcrest "E" oil, 525 SUV at 100 F; 98 cp at 100 F, and 28 cp at 150 F.)

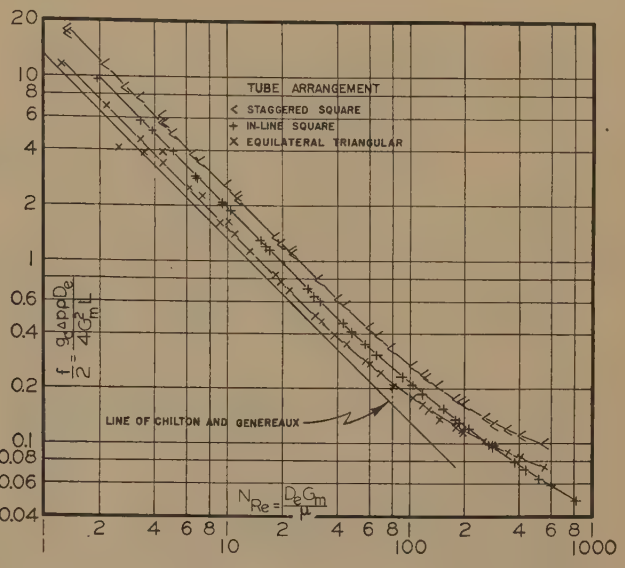


FIG. 4 FRICTION FACTOR VERSUS REYNOLDS NUMBER UNDER ISOTHERMAL CONDITIONS FOR THREE TUBE ARRANGEMENTS

(Flow normal to 3/8-in-OD vertical tubes, 15/32-in. pitch. Average bulk-oil temperatures 125°, 150°, and 175°F. Gulfcrest "E" oil, 525 SUV at 100 F; 98 cp at 100 F, 49 cp at 125 F, 28 cp at 150 F, and 17 cp at 175 F. Note: The line of Chilton and Genereaux is based upon data from two equilateral triangular arrangements with 1.25 and 1.59 pitch-to-diameter ratios using 3/4-in. tubes.)

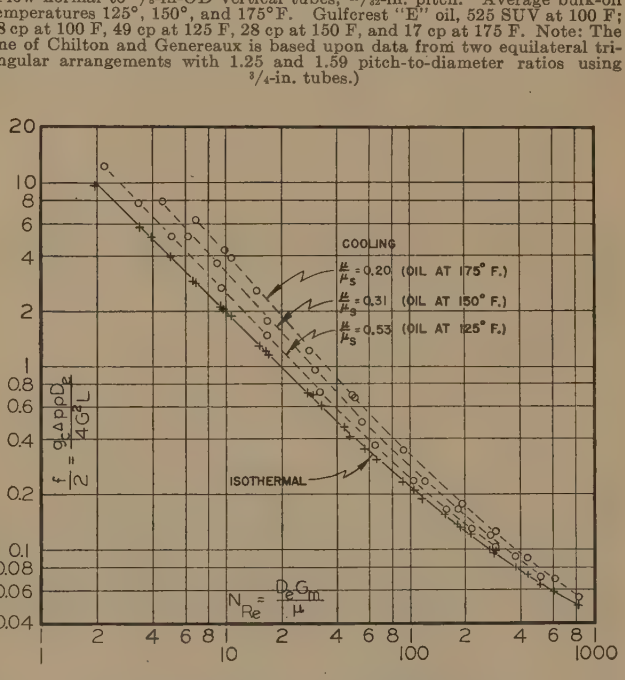


FIG. 5 FRICTION FACTOR VERSUS REYNOLDS NUMBER FOR IN-LINE SQUARE TUBE ARRANGEMENT

(Flow normal to 3/8-in-OD vertical tubes, 15/32-in. pitch. Average bulk-oil temperature for cooling runs, 125, 150, and 175 F; surface-temperature range for cooling runs, 100–109 F; average bulk-oil temperature for isothermal runs, 125, 150, and 175 F. Gulfcrest "E" oil, 525 SUV at 100 F; 98 cp at 100 F, 49 cp at 125 F, 28 cp at 150 F, and 17 cp at 175 F.)

the transition region indicate that it would be difficult to obtain a correlating function to cover the entire range of flow.

While it has been shown that the more complicated method of Gunter and Shaw (5) is about as satisfactory as the foregoing method for representing data from this type of exchanger, the simpler method was chosen for use at this time because it represents the data nearly as well and also because not all of the variables in the Gunter and Shaw correlation have been studied. Later, when these factors have been treated experimentally, a critical comparison of the two methods can be made or a new correlation presented.

Another variation of fluid properties occurs when heat transfer causes the physical properties of the fluid at the tube wall to differ from those within the body of the fluid. The effect of this variation is shown in Fig. 5 for the in-line square arrangement. The isothermal data at levels of 125, 150, and 175 F all fall along the solid line, but cooling data, taken with the tube wall at approximately 100 F, and the bulk-oil temperature at average values of 125, 150, and 175 F, are shown, respectively, by higher levels of the dashed lines. The degree of increase in pressure drop appears to be a function of the ratio of the bulk-oil viscosity to the viscosity at the tube wall, and also an inverse function of the Reynolds number.

When the pressure-drop data in Fig. 5 are corrected by the addition of a term in the friction factor to compensate for the effect of the viscosity gradient between the tube wall and the body of the oil stream, a single line is obtained for each tube arrangement as shown in Fig. 6. The correction term is the viscosity ratio (μ/μ_s) raised to the power n, where n is a function of the Reynolds number having the form $n = 0.56 - 0.16 \log N_{Re}$. Values for n are plotted versus the Reynolds number at the bottom of Fig. 6. The function relating n to the Reynolds number is obtained by calculating the values of n required to bring the friction factors for cooling into agreement with the isothermal friction factors, plotting the calculated values versus the Reynolds number, and then drawing the best line through these values. The details of this operation are reported in the forthcoming research bulletin (3). It is apparent that an excellent correction for the effect of the viscosity difference is obtained, but it should be emphasized that for the present the correction applies only to the current data for the in-line square arrangement. For the case of viscous flow inside tubes, Sieder and Tate (6) found that the function $1.1(\mu/\mu_s)^{0.25}$ correlated friction data under nonisothermal conditions. This function does not correlate the current data satisfactorily, because it is independent of the Reynolds number. It is possible that a part of the present deviation at low linear velocity may be due to nonuniformity of flow and free convection. The extent of these effects can be determined most satisfactorily by the use of more viscous oils and various tube spacings and will be investigated later in the program. It is believed that the present

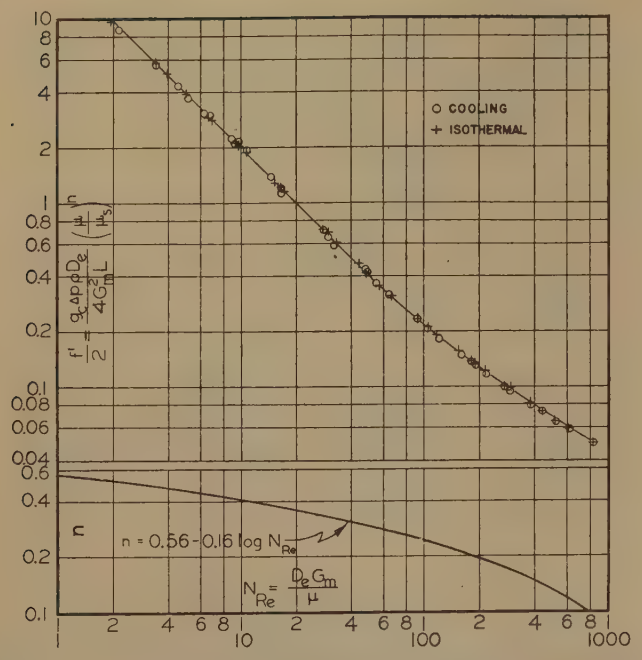


FIG. 6 CORRECTED FRICTION FACTOR VERSUS REYNOLDS NUMBER FOR IN-LINE SQUARE TUBE ARRANGEMENT

(Flow normal to 3/8-in-OD vertical tubes, 15/32-in. pitch. Average bulk-oil temperature for cooling runs, 125, 150, and 175 F; surface-temperature range for cooling runs, 100–109 F; average bulk-oil temperature for isothermal runs, 125, 150, and 175 F. Gulfcrest "E" oil, 525 SUV at 100 F; 98 cp at 100 F, 49 cp at 125 F, 28 cp at 150 F, and 17 cp at 175 F.)

development represents as much as can be learned from one oil and one tube spacing. Similar correcting functions have been developed for the equilateral and staggered square arrangements, but the search for a single function which will be adequate for various arrangements as a part of a general correlation will be postponed until other units and oils have been studied.

Heat Transfer. The experimental data on heat transfer are shown in Fig. 7 where the oil-film coefficient of heat transfer is plotted versus the velocity through the minimum area for flow between tubes. The data from the triangular and staggered square arrangements can be represented by a single line, while points for the in-line square arrangement fall about 40 per cent below the staggered-tube data at the lowest velocity, and about 20 per cent below at the highest velocity. The increase in slope at velocities of about 1 fps is believed to be due to the start of turbulence. The data in Fig. 7 are for an average oil temperature of 150 F. Data at other temperature levels give similar but

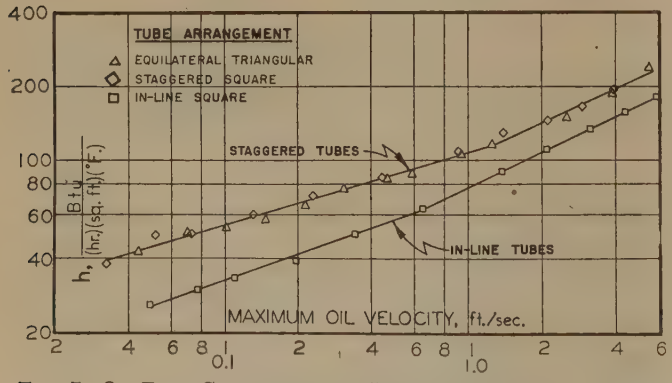


FIG. 7 OIL-FILM COEFFICIENT OF HEAT TRANSFER FOR THREE TUBE ARRANGEMENTS

(Bulk-oil temperature range 140–160 F; surface-temperature range 100–109 F. Flow normal to 3/8-in-OD vertical tubes, 15/32-in. pitch. Gulfcrest "E" oil 525 SUV at 100 F; 98 cp at 100 F, and 28 cp at 150 F.)

displaced lines, since this method of representation makes no correction for changes in the physical properties of the fluid.

A more general correlation is given in Fig. 8 where the j factor

$$\frac{h}{cG_m}\left(\frac{c\mu}{k}\right)^{2/3}\left(\frac{\mu_s}{\mu}\right)^{0.14}$$

is plotted versus the Reynolds number, D_eG_m/μ. The data for each arrangement fall along separate lines, and as yet no factor has been obtained which will correlate the data for all three units. An equivalent diameter, based upon the geometry of tube arrangement, shows some promise, but additional data are needed to develop the correlation. The data in Fig. 8 cover bulk-oil tempera-

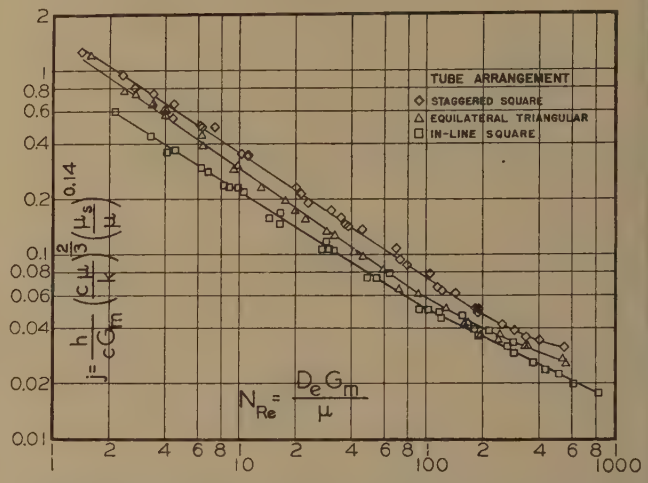


FIG. 8 . HEAT-TRANSFER FACTOR j FOR THREE TUBE ARRANGEMENTS

(Flow normal to 3/8-in-OD vertical tubes, 15/32-in. pitch. Average bulk-oil temperatures 125, 150, and 175 F; surface-temperature range 100–109 F. Gulfcrest "E" oil 525 SUV at 100 F; 98 cp at 100 F, 49 cp at 125 F, 28 cp at 150 F, and 17 cp at 175 F.)

tures of 125, 150, and 175 F, and a tube-wall temperature of approximately 100 F, which gives μ_s/μ ratios of approximately 1.9, 3.2, and 5.0, respectively. The data are correlated fairly well over the entire range of Reynolds numbers, by the single exponent of 0.14 for μ_s/μ, similar to the value found for flow inside tubes by Sieder and Tate (6). The single exponent is in striking contrast to the wide variation of exponent necessary for a satisfactory correlation of pressure drop.

Fig. 8 does not necessarily present a true comparison of the effectiveness of heat transfer for the three tube arrangements, because a different choice of terms in the Reynolds number will change the relative position of the lines. A better comparison is given in Fig. 9 where the heat-transfer coefficient is shown versus the power loss per unit heat-transfer area. On this basis the data for the two staggered arrangements fall together and well above those for the in-line square arrangement. It can be seen that in the viscous region the staggered arrangements give superior heat transfer for a given power loss, but the superiority becomes less marked as the turbulent region is approached. Fig. 9 applies only to viscous flow, but it is interesting to note that for turbulent flow, using a similar method of comparison and the data of Pierson, et al. (1) for gases, in-line arrangements were shown by Colburn (7) to have a slight superiority over staggered arrangements. The apparent convergence of the data in the transition region, as shown in Fig. 9, may foretell this inversion, but data on the same tube banks in both viscous and turbulent flow are needed to clarify the point.

From the tests on these three simple tube banks it appears that, in the region of viscous flow, for the same power loss per unit surface area, the heat-transfer coefficients for staggered tubes are

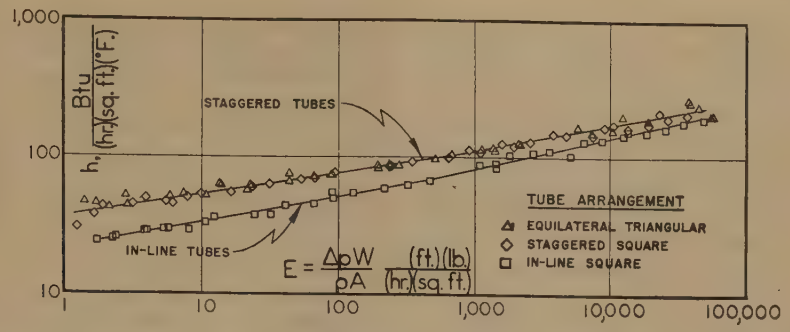


FIG. 9 COMPARISON OF HEAT-TRANSFER AND POWER REQUIREMENTS FOR THREE TUBE ARRANGEMENTS

(Flow normal to 3/8-in-OD vertical tubes, 15/32-in. pitch. Bulk-oil temperature range 115–200 F; surface-temperature range 100–109 F. Reynolds number range 2-800. Gulfcrest "E" oil 525 SUV at 100 F; 98 cp at 100 F, 49 cp at 125 F, 28 cp at 150 F, and 17 cp at 175 F.)

some 50 per cent higher than for in-line tubes, and, at the same heat-transfer coefficient, the pumping power for the in-line arrangement is around tenfold that for staggered arrangements. Caution must be used, however, in interpreting these results in terms of commercial shell and tube exchangers where baffle spacing, baffle cutoff, economic fluid velocity, and the diameter of the exchanger must be considered. For example, whereas the staggered square spacing gave about the same performance as triangular and might be preferred because of greater ease of cleaning, it should be noted that for the same velocity through minimum area, a closer baffle spacing would be required. Quantitative information concerning the effect of baffles is needed to be able to utilize the present results for the design of baffled heat exchangers.

CONCLUSIONS

1 In the region of viscous flow and for a pitch-to-diameter ratio of 1.25, equilateral-triangular and staggered square-tube arrangements are approximately equal in effectiveness for heat transfer. The in-line square arrangement has a lower degree of effectiveness in viscous flow, but its performance approaches that of the staggered arrangements in the transition and early turbulent regions.

2 Viscosity-ratio functions, derived from the experimental data, are effective in bringing the isothermal and nonisothermal fluid-friction data into agreement. The Sieder and Tate viscosity-ratio correction is adequate for the heat-transfer data.

3 In correlating data in the region of viscous flow, the use of the equivalent diameter and the mass velocity through the minimum cross section is insufficient to compensate for the effect of differences between tube patterns.

ACKNOWLEDGMENT

The authors acknowledge the suggestions on design of the apparatus and on the program of study by the Special Advisory Committee of the Heat Transfer Division, which is sponsoring this co-operative research project. The committee at the time of the work was composed of W. H. McAdams, Chairman, S. Kopp, A. C. Mueller, B. E. Short, W. H. Thompson, T. Tinker, and P. R. Trumpler. The help of A. Wurster in the design and construction of the test units was invaluable. Assistance in interpretation and presentation of data was given by A. P. Colburn; Research Fellows W. E. Meece and F. W. Sullivan aided in the experimental work. Funds and equipment were kindly furnished by the Tubular Exchanger Manufacturer's Association, The American Petroleum Institute, Andale Company, Davis Engineering Company, Downingtown Iron Works, du Pont Company, Standard Oil Development Company, and York Ice Machinery Company. The oil was provided by the Gulf Oil Company.

BIBLIOGRAPHY

1 "Experimental Investigation of the Influence of Tube Arrangement on Convection Heat Transfer and Flow Resistance in Cross Flow of Gases Over Tube Banks," by O. L. Pierson, Trans. ASME, vol. 59, 1937, pp. 563–572. See also papers by E. C. Huge, pp. 573–581, and E. D. Grimison, pp. 583–594, and discussion, vol. 60, 1938, pp. 381–391.

2 "Heat Transfer and Fluid Friction During Viscous Flow Across Banks of Tubes," by G. A. Omohundro, O. P. Bergelin, and A. P. Colburn, Trans. ASME, vol. 71, 1949, pp. 27–34.

3 "Heat Transfer and Fluid Friction During Flow Across Banks of Tubes," by O. P. Bergelin, A. P. Colburn, and H. L. Hull, University of Delaware Engineering Experimental Station Bulletin No. 2, 1949.

4 "Pressure Drop Across Banks of Tubes," by T. H. Chilton and R. P. Genereaux, Trans. AIChE, vol. 29, 1933, pp. 161–173.

5 "A General Correlation of Friction Factors for Various Types of Surfaces in Crossflow," by A. Y. Gunter and W. A. Shaw, Trans. ASME, vol. 67, 1945, pp. 643–660.

6 "Heat Transfer and Pressure Drop of Liquids in Tubes," by E. N. Sieder and G. E. Tate, *Industrial and Engineering Chemistry*, vol. 28, 1936, pp. 1429–1435.

7 "Heat Transfer by Natural and Forced Convection," by A. P. Colburn, Purdue University Experimental Station, Bulletin No. 84, 1942.

Discussion

C. H. BROOKS.[5] It is to be hoped that the authors will endeavor to put their correlations in such form that the designer will have a minimum of trial-and-error calculation to make when designing an exchanger. In general, most of the equations which have been published are fine correlations of experimental data and are useful for determining whether test data approach theory closely or not, but difficult for the average engineer to use when designing a new piece of equipment. This is the case because so many of the variables in the published relations need to be assumed and are dependent upon other design variables which may be affected by the heat-transfer coefficients resulting from the calculation.

A. S. FOUST.[6] The most significant point upon which the writer would like to request discussion by the authors is the difference in the shapes of the friction-factor and j-factor plots at Reynolds numbers in the transition region. It appears difficult to explain the reversals in the friction-factor plots which appear at Reynolds numbers somewhat over 100 in contrast to the uniform shape maintained by the j-factor plots to the maximum Reynolds number studied.

It will be highly illuminating to the profession, when this en-

[5] Process Engineer, Sun Oil Company, Philadelphia, Pa.

[6] Professor, Department of Chemical and Metallurgical Engineering, University of Michigan, Ann Arbor, Mich.

tire program is completed, to verify the use of an equivalent diameter and maximum mass velocity in the Reynolds number. This question arises in the mind of the writer in view of the difference in Reynolds numbers so computed at which the friction factors indicate the incipient breakdown of laminar flow. It is noted that the transition begins earliest in the "in line square" arrangement at a Reynolds number under 100, develops next in the staggered square arrangement at a Reynolds number approximately 200, and is barely perceptible in the equilateral-triangle arrangement at the maximum Reynolds number studied.

ALFRED GERMAN.[7] The experiments now in progress by the authors are excellent and constitute a step in the right direction. Operating conditions, however, as a rule do not permit a duplication of results as obtained by such experiments. Oil-refinery conditions usually develop considerable fouling on the shell side where elevated temperatures are involved. In the experiments described, a clean medium is used to flow around the tubes placed in either triangular or square pitch layout in order to indicate the most desirable pattern. Actual practical experience has shown that, particularly in the case of the triangular-tube layout, fouling is in many cases so rapid that, if allowed to continue, the tube bundle would be completely filled in a very short time. We have known cases where the entire shell side of the unit has been completely filled within 2 weeks.

When changes were made to vertical laning in a horizontal unit, a decided improvement resulted. This tube bundle, under identical operating conditions, could continue in service several times longer than could the triangular-tube layout.

Even though the vertical lanes with the help of up and down crossflow baffling keep the tube bundle cleaner for a longer time of operation, there is still further room for improvement. Using this vertical laning in the large-diameter tube bundle with the spaces between the tubes only $^3/_{16}$ or $^1/_4$ in. there is still considerable accumulation of solids between the tubes.

This accumulation can be diminished by increasing the tube pitch. However, this expedient will result in a reduction in surface due to the smaller number of tubes which can be installed in a shell of a given diameter. The writer believes that the initial heat-transfer loss due to the smaller area will be more than compensated for by the relative freedom from fouling, thus resulting in a better average over-all heat transfer over a period of time.

It would be interesting to repeat these experiments, using a medium calculated to simulate actual fouling conditions normally encountered in extremely severe service. For example, instead of a clean medium, it is suggested that a viscous liquid containing fine particles, heavier than the liquid, which will tend to deposit on the tubes, may duplicate to some extent the fouling under consideration.

In this way information might be obtained which would indicate not only the relative desirability of various tube layouts but also, in some degree, the minimum spacing for various types of media and for various bundle diameters.

E. N. SIEDER.[8] The authors have presented interesting and valuable data which the heat-exchanger industry has been awaiting for a long time. We hope that the data for turbulent flow, which we understand are being obtained on the tube bundles described, when correlated, will clear up once and for all the important question, "which gives the best heat transfer for a given pressure drop-tubes on a staggered square pitch or tubes on an in-line square pitch?"

The authors have shown that for viscous flow the staggered-tube arrangement is apparently the best. They point out that this condition may not hold true in the turbulent region.

A correlation of heat-transfer data in turbulent flow, which was prepared by Gunter and Shaw but which was never published, indicates that the square in-line tube arrangement is superior to the square staggered arrangement in so far as heat-transfer rates for a given pressure drop are concerned. Their data show that, for a given baffle pitch, the heat-transfer rate obtainable with square in-line pitch is 43 per cent better than the rate obtained with staggered pitch, while the pressure drop obtained is 2.9 times as great. When the baffle pitch is adjusted to give equal pressure drops, the heat-transfer rate for the square in-line arrangement is 13 per cent better than for the staggered square pattern.

Moreover, to obtain the same pressure drop with a square staggered pitch, 45 per cent more baffles have to be used in a given exchanger than are required for a square in-line pitch.

These figures are quoted to show that the question is far from being solved, and that we may never be able to show a clear-cut advantage for either case.

AUTHORS' CLOSURE

Dr. Brooks's suggestion that correlations be presented in a form suitable for use in design is in accord with present plans. It is felt, however, that such correlation can be made more satisfactorily at a later date when data covering a wider range of variables are available. For design purposes, furthermore, consideration must be given to the practical problems of baffle and tube spacings, leakage, etc.

With regard to the points raised by Professor Foust, it is difficult to determine the exact point where the change from viscous to turbulent flow occurs because the change is gradual and covers a range of Reynolds numbers. The question of what equivalent diameter and velocity to use should be clarified when data have been obtained with other arrangements and with oils having other viscosities.

The deviations caused by fouling are of primary importance in the practical use of these data but a study of fouling is not a part of the immediate program. It is hoped that the data on nonfouling fluids may serve as the basis for later studies under fouling conditions. The approach suggested by Mr. German seems sound and might be used with the present models and auxiliary equipment at a later date.

An extension of the present tests into the region of turbulent flow with both in-line and staggered square arrangements is a part of the present plans. It is hoped that some of the points raised by Mr. Sieder may be clarified by this study, although again the complications in a baffled heat exchanger make general conclusions difficult to draw from data on unbaffled units.

[7] Technical Service Engineer, Heat Exchanger Tubes, Scovill Manufacturing Company, Inc., New York, N. Y.

[8] District Sales Manager, Alco Products Division, American Locomotive Company, New York, N. Y.

Investigation of Variation of Point Unit Heat-Transfer Coefficient Around a Cylinder Normal to an Air Stream

By W. H. GIEDT,[1] BERKELEY, CALIF.

Most previous analytical and experimental work concerned with the variation of the rate of heat transfer around a cylinder normal to an air stream has been confined to one with an isothermal surface. It was therefore the purpose of the investigation described in this paper to obtain information on the variation of the rate of heat transfer around a cylinder with a nonisothermal surface. The method was suggested by a technique used in flight tests for determining point unit heat-transfer coefficients around airfoils. A thin nichrome ribbon wound helically around a cylinder was electrically heated. Point unit heat-transfer coefficients were determined from the temperature variation along, and electrical input to the ribbon. Results indicated the method would also be suitable for similar measurements around other streamlined shapes. The experimental data include point unit heat-transfer coefficient and static-pressure distributions around a cylinder circumference throughout a range of Reynolds modulus from 70,800 to 219,000. Comparison of these results with the skin-friction distributions around the cylinder indicated by the pressure distributions showed a definite correlation between the two. Results are also compared with those obtained by other investigators using an isothermal cylinder surface.

Nomenclature

The following nomenclature is used in this paper:

A = area of heat-transfer surface, sq ft
D = outside diameter of cylinder with nichrome ribbon around it, ft
h = unit surface heat-transfer coefficient, Btu/hr ft² deg F
h_θ = unit heat-transfer coefficient at cylinder surface at angle θ, Btu/hr ft² deg F
i = current flowing through nichrome ribbon, amp
k_a = thermal conductivity of air, Btu/hr ft² (deg F/ft)
k_n = thermal conductivity of nichrome ribbon, Btu/hr ft² (deg F/ft)
p = static pressure at any point on the cylinder surface, psf
p_0 = free-stream static pressure, psf
q = rate of heat transfer by convection from cylinder surface at any point defined by the angle θ, Btu per hr
q_{cond} = rate of heat loss from nichrome ribbon through cylinder wall, Btu per hr
R = electrical resistance of nichrome ribbon at 70 F, ohms per ft of length
t = temperature of nichrome ribbon at any point around the cylinder, deg F
t_{min} = minimum temperature of nichrome ribbon for a given power input and free air-stream velocity, deg F
t_{max} = maximum temperature of nichrome ribbon for a given power input and free air-stream velocity, deg F
t_0 = free air-stream temperature, deg F
$t_{avg} = t_0 + \frac{t_{max} - t_{min}}{2}$ = average temperature at which properties of air were evaluated, deg F
T = absolute temperature, deg R
u_0 = free-stream velocity, fps
x = distance around cylinder circumference measured from stagnation point, ft
y = distance from nichrome ribbon measured perpendicular to cylinder surface, ft
ϵ = total hemispherical emissivity
μ = absolute viscosity of air, lb-sec/ft²
θ = angle at center of cylinder measured from the forward stagnation point, deg
ρ_0 = mass density of air in free stream, lb-sec²/ft⁴
σ = Stefan-Boltzmann constant, 0.1782 Btu/hr (deg R/100)⁴ft²
τ_0 = shear stress at wall of cylinder, psf
$\tau_0/\rho_0 u_0^2$ = intensity of skin friction at cylinder wall
$(N_{Nu})_\theta = \frac{h_\theta D}{k_a}$ = Nusselt modulus based on h_θ
$N_{Re} = \frac{D\, u_0\, \rho_0}{\mu}$ = Reynolds modulus based on cylinder diameter

Introduction

It has long been recognized that a knowledge of the peripheral variation of the rate of heat transfer from the surface of streamlined shapes to a fluid flowing transversely is necessary for the successful design of heat-transfer equipment. Basic to this problem is the heat transfer from a right circular cylinder, and considerable interest has been exhibited in it since an early date. Experimental investigations have led to the development of expressions describing the variation of the average heat-transfer coefficient (1)[2] and the variation of the local heat-transfer coefficient from the stagnation point back to approximately the 80-deg point (2, 3). However, because of the complexity of the problem up to the present time, no satisfactory complete analytical solution has been found.

Experimental investigations of the rate of heat transfer from a cylinder have not been as extensive as might be desired, and

[1] Instructor, Division of Mechanical Engineering, University of California.

Contributed by the Heat Transfer Division and presented at the Heat Transfer and Fluid Mechanics Institute, Los Angeles, Calif., June 21, 1948, of The American Society of Mechanical Engineers.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

have, in general, been restricted to cylinders with an isothermal surface. It is realized, however, that most heat transfer actually takes place from a nonisothermal cylinder surface. In view of this and the fact that practically no experimental data are available for this case, it was the purpose of the experiment described in this paper to investigate the variation of the point heat-transfer coefficient around a cylinder with a nonisothermal surface.

The method used was to wrap a thin metal ribbon around a cylinder parallel to the air stream. An electric current passed through the ribbon generated equal amounts of energy in each differential element of the ribbon. If these equal amounts of heat are transferred to the air stream according to the relation,

$$\frac{q}{A} = h_\theta \Delta t$$

where Δt = the difference between the ribbon temperature and the free-air temperature, it is seen that a variation in h_θ will cause a variation in Δt. Since $(q/A) \sim i^2R$, h_θ could be calculated from

$$h_\theta = \frac{(q/A)}{\Delta t} \sim \frac{i^2R}{\Delta t}$$

Experimental Equipment

The test cylinder is shown in Fig. 1. It was mounted for testing in the University of California 3-ft wind tunnel as shown in Fig. 2.

The cylinder itself was made in three sections from 4-in-OD lucite stock. A length of 40-in. gave it a span greater than the wind-tunnel air stream, thereby providing two-dimensional flow. A nichrome ribbon 1 in. × 0.002 in. in cross section, wound helically around the center section of the cylinder, formed the heating element. A length of 65 in. provided for five full turns of the ribbon, the center being the test strip, and the two on either side acting as guard heating strips. Both ends of the ribbon were led through the cylinder wall where they were connected to power leads brought in through one end of the cylinder. The power supply to the ribbon was controlled through a constant-voltage transformer which held voltage variation with time to within 1 per cent.

The temperature variation along the ribbon was measured by means of very fine thermocouples. Short lengths of B & S gage No. 40 iron and constantan wire were spot-welded to the

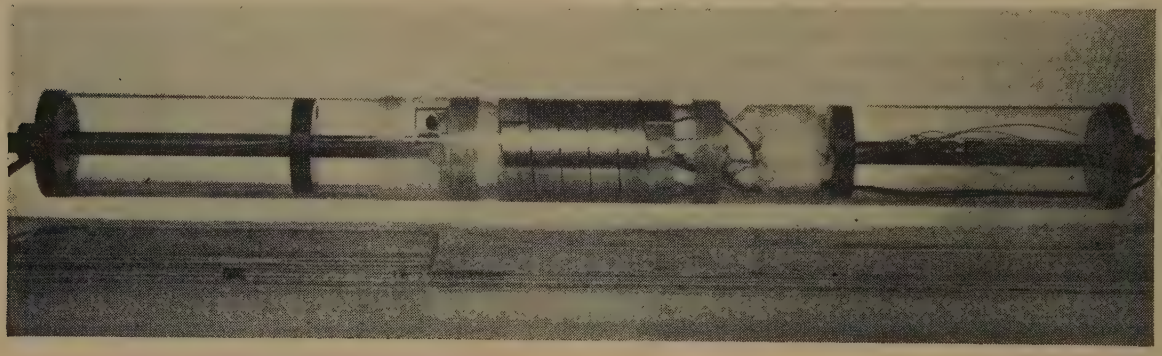


Fig. 1 Lucite Test Cylinder With Nichrome-Ribbon Heating Element and Pressure Tap

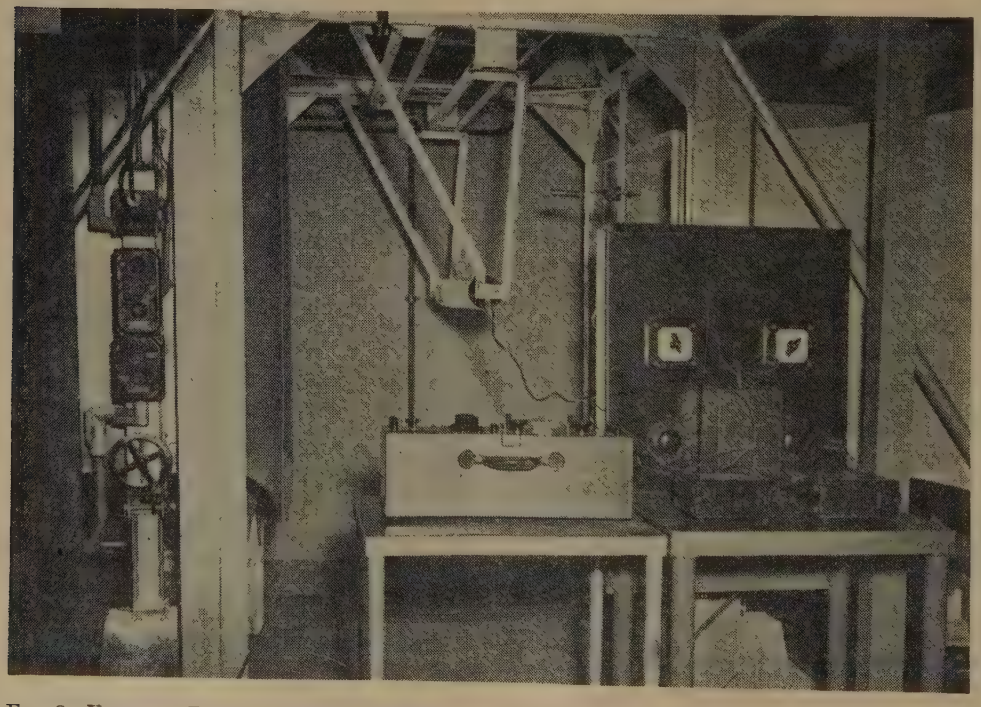


Fig. 2 View of Test Cylinder Mounted in University of California 3-Ft Wind Tunnel With Test Panel and Instruments Used for Taking Data

nichrome ribbon before it was placed around the cylinder. The thermocouples so formed were located so that they would be spaced about $^1/_2$ in. apart for one half of the circumference on the center turn of the ribbon when wrapped around the cylinder, as shown in Fig. 1. These short lengths of wire were led through the cylinder wall and soldered to leads brought in through one end of the cylinder. Before final assembly the center section of the cylinder was filled with glass wool to reduce conduction losses.

A static-pressure tap, located in the cylinder wall near the center section, Fig. 1, provided the means of determining circumferential static-pressure variation by rotating the cylinder. Free air temperature was measured with a shielded thermocouple of B & S gage No. 30 iron and constantan wire.

Experimental Procedure

The shielded thermocouple used for measuring the free-air-stream temperature was first calibrated. Then, because only temperature differences between the nichrome ribbon and the free air stream were involved in the calculation of the h_θ values, the ribbon thermocouples were calibrated against the free-air-stream thermocouple. The electrical resistance of the nichrome ribbon was determined from measurements made on a 10-ft length at 70 F with a Wheatstone bridge, the accuracy of which was $^1/_{20}$ per cent. The value of 0.262 ohm per ft obtained was considered constant in the calculations in so far as the maximum variation of the resistance throughout the experimental testing was less than 0.5 per cent.

The amount of heat lost by the nichrome ribbon through the cylinder wall was determined by wrapping a thick layer of glass-wool insulation around the section covered by the ribbon. Electrical power inputs necessary to maintain the nichrome ribbon at different temperatures above the free air temperature and simultaneous temperature distributions in the glass wool were measured. Losses through the steel wool were calculated. Subtracting these from the ribbon inputs gave the heat loss through the cylinder walls as a function of ribbon temperature. These amounted to a maximum of 2 per cent of ribbon electrical input during the experimental runs.

Nichrome-ribbon temperature-distribution and static-pressure-distribution data around the cylinder circumference were taken throughout a Reynolds modulus range of 70,000 to 219,000. In general, measurements were limited to the top half of the cylinder circumference; however, as a check on the symmetry, several were made around the entire cylinder circumference, without any substantial differences between top and bottom being manifested.

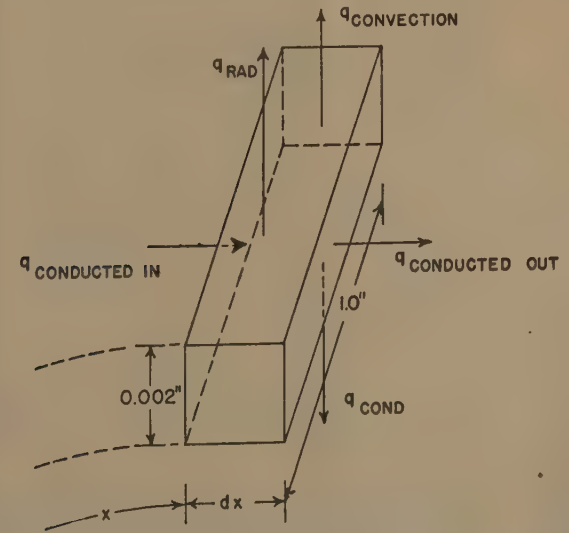


Fig. 3 Heat-Flow Diagram for a Differential Length of Nichrome Ribbon

Calculation of Point Unit Heat-Transfer Coefficients

Considering the differential length of nichrome ribbon shown in Fig. 3 under steady-state conditions yields the following heat balance

$$\left.\begin{aligned} &\begin{pmatrix}\text{heat generated}\\ \text{by current } i\end{pmatrix} + \begin{pmatrix}\text{heat conducted}\\ \text{in left face}\end{pmatrix} = \\ &\begin{pmatrix}\text{heat conducted}\\ \text{out right face}\end{pmatrix} = \begin{pmatrix}\text{heat convected}\\ \text{to air stream}\end{pmatrix} + \\ &\begin{pmatrix}\text{heat radiated}\\ \text{to surroundings}\end{pmatrix} + \begin{pmatrix}\text{heat lost by conduction}\\ \text{through cylinder wall}\end{pmatrix} \end{aligned}\right\} \ldots [1]$$

Or in symbolic form

$$i^2R\,dx + \left[-k_n \frac{1 \times 0.002}{144}\left(\frac{dt}{dx}\right)\right] + \left[\left\{-k_n \frac{1 \times 0.002}{144}\right\} \left\{\frac{dt}{dx} + \frac{d}{dx}\left(\frac{dt}{dx}\right)\right\} dx\right]$$

$$+ h_\theta \frac{1}{12}\,dx\,(t - t_0) + \frac{q_{rad}}{A}\frac{1}{12}\,dx + \frac{q_{cond}}{A}\frac{1}{12}\,dx \ldots\ldots [2]$$

Substituting $k_n = 7.86$ Btu/hr ft^2 (deg F/ft), $R = 0.262$ ohm per ft, expressing the distance around the cylinder in terms of θ, and solving for h_θ gives

$$h_\theta = \frac{10.73i^2 + 154\left(\frac{d^2t}{d\theta^2}\right) - \frac{q_{rad}}{A} - \frac{q_{cond}}{A}}{(t - t_0)}$$

Values of i were measured. The radiation loss from each point was calculated from

$$\frac{q_{rad}}{A} = \epsilon\sigma\left[\left(\frac{T}{100}\right)^4 - \left(\frac{T_0}{100}\right)^4\right]$$

using a value of 0.2 for ϵ (ref. 8); $\frac{q_{cond}}{A}$ for each point was determined as described in the previous section. Since the temperature variation around the cylinder did not follow any known law, it was necessary to evaluate the $(d^2t)/(d\theta^2)$ term at each point graphically. For each experimental run the temperature potential of the nichrome ribbon above the free air stream was plotted versus θ (Fig. 4 is a representative curve). Curves of $(dt)/(d\theta)$ versus θ (see Fig. 5), were then plotted from values of $(dt/d\theta)$, determined graphically from the temperature-potential curves. Values of $(d^2t/d\theta^2)$ were then determined from these curves. These h_θ values multiplied by the cylinder diameter and divided by k_a, evaluated at an average film temperature, gave the Nusselt modulus at each point.

It may be of interest to note that the $(d^2t)/(d\theta^2)$ term represents the effect of heat conduction along the ribbon. Its effect on h_θ, in general, amounted to only a few per cent except near the maximum and minimum points, where it was as much as 15 per cent.

Heat-Transfer and Pressure Distributions

It has been found that the point unit heat-transfer coefficient, h_θ, for a cylinder losing heat to a normal air stream depends on the Reynolds modulus, N_{Re}, of the flow and on the location along the cylinder circumference. Introducing the point unit heat-

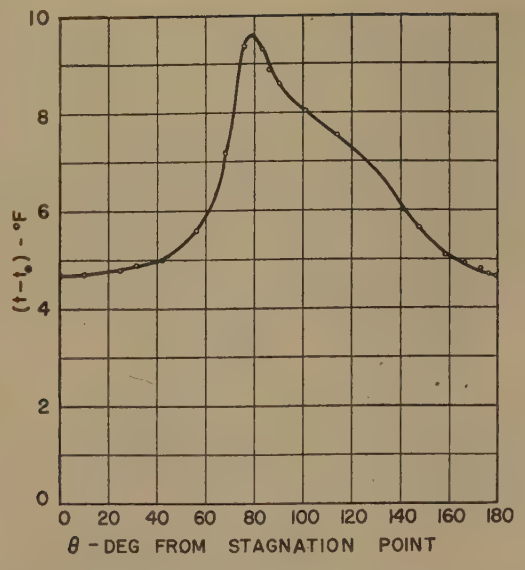


FIG. 4 TEMPERATURE DISTRIBUTION OF NICHROME RIBBON AROUND CIRCUMFERENCE OF THE CYLINDER PLOTTED IN TERMS OF INCREMENT, $t - t_0$, ABOVE FREE-STREAM TEMPERATURE (Reynolds modulus = 101,300.)

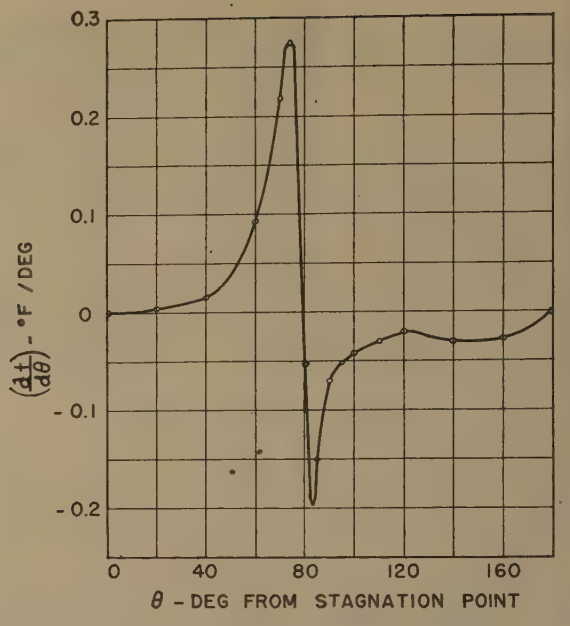


FIG. 5 RATE OF CHANGE OF TEMPERATURE OF NICHROME RIBBON WITH RESPECT TO ANGLE AROUND CYLINDER CIRCUMFERENCE FOR REYNOLDS MODULUS = 101,300

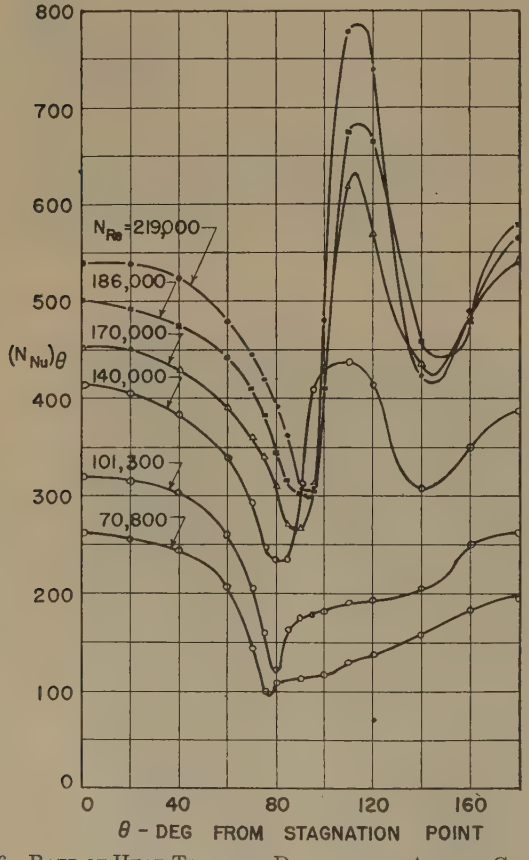


FIG. 6 RATE OF HEAT-TRANSFER DISTRIBUTION AROUND CYLINDER CIRCUMFERENCE FOR DIFFERENT VALUES OF REYNOLDS MODULUS; $h_\theta = 0.044\ (N_{Nu})_\theta$

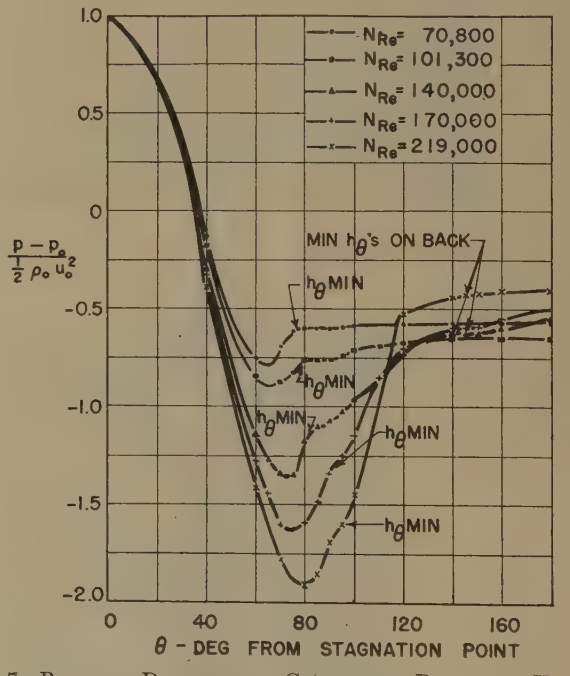


FIG. 7 PRESSURE-DISTRIBUTION CURVES FOR DIFFERENT VALUES OF REYNOLDS MODULUS

transfer coefficient in nondimensional form as the Nusselt modulus

$$(N_{Nu})_\theta = \frac{h_\theta D}{k_a}$$

the following relation can be written

$$(N_{Nu})_\theta = F(N_{Re}, \theta)$$

where θ characterizes the location of h_θ around the cylinder circumference in terms of the angle measured from the forward stagnation point.

The values of $(N_{Nu})_\theta$ calculated from the experimental data are plotted against θ in Fig. 6 for the different values of Reynolds modulus. The pressure-distribution measurements are similarly plotted in Fig. 7. The following points can be observed in these plots:

1 Maximum values of h_θ occur approximately 10 deg further back along the cylinder circumference than the minimum static-presssure values.

2 Both minimum h_θ values and pressures move back along the cylinder circumference with increasing values of N_{Re}.

3 For values of N_{Re} up to around 100,000 maximum h_θ values occur at the 0-deg and 180-deg points.

4 At values of N_{Re} = 140,000 and up h_θ at the rear stagnation point is greater than at the forward stagnation point.

5 At values of N_{Re} from 140,000 to 219,000 distinct maximum h_θ values occur between 110 deg and 115 deg.

6 For this same range of N_{Re} two distinct minimum h_θ values occur. The lower of the two is between 80 and 93 deg on the front half of the cylinder. The second is between 140 and 150 deg on the back half.

7 Also, for this same range of N_{Re}, inflection points occur in the pressure distribution curves at the same locations as the minimum h_θ values just noted.

Correlation of these results with the results of investigations of the air flow in the boundary layer around a circular cylinder made by Fage and Falkner (4, 5) brings out some interesting parallels. In brief, their static-pressure and skin-friction distribution measurements indicated the following:

For flow in the critical range of N_{Re}, i.e., the range where large changes in the cylinder drag coefficient occur, there was a critical point on the cylinder, just beyond the region of maximum negative pressure, where a transition from laminar to turbulent flow in the boundary layer began. The position of this point on the cylinder circumference was marked by a pronounced inflection in the static-pressure curve. The skin-friction measurements showed two types of distributions around a cylinder circumference. The first type, which occurred at relatively low values of N_{Re}, rose gradually from zero at the stagnation point to a maximum around θ = 55 deg, and then fell rapidly to zero. The second type, which occurred in the critical region of N_{Re}, after reaching a maximum around θ = 55 deg fell less rapidly to a minimum value around θ = 80 deg and then rose to a second maximum before dropping to zero around 130 deg.

They concluded that the point where the frictional intensity falls to its zero value indicates the position on the cylinder where separation of the boundary layer from the surface is completed. For values of N_{Re} below the critical range, this apparently occurred before the 90-deg point. For values of N_{Re} in the critical range, the minimum frictional intensity around 90 deg was considered to indicate the position along the surface where flow in the boundary layer changed from laminar to turbulent; separation of the boundary layer was considered to occur at the point on the back half of the cylinder where the frictional intensity reached its zero value. At the lower values of N_{Re} the pressure distribution was noted to level off to a constant value where the frictional intensity reached zero. In the critical range a pronounced inflection in the pressure distribution occurred at the point where the frictional intensity reached a minimum; the point where the frictional intensity dropped to zero was indicated by a leveling-off of the pressure curve. These investigators also found that increasing the turbulence in the free air stream caused static-pressure and frictional-intensity distributions characteristic of the critical range to occur at values of N_{Re} below that range.

If Figs. 6 and 7 are now interpreted in light of this information, the following can be noted:

1 For the run with N_{Re} = 70,800, the skin-friction distribution was probably of the first type just described, in so far as the pressure-distribution curve, after passing through a minimum, levels off at a constant value around 78 deg. This point then marks the location at which the skin friction dropped to zero and at which the flow separated from the surface of the cylinder. The rate of heat transfer would be expected to be low at this point, which is verified by the fact that the location of the minimum h_θ coincides with it.

2 For the run with N_{Re} = 101,300 the minimum h_θ occurs at the location where the pressure distribution levels off to a constant value for a distance corresponding to about 10 deg. The pressure rises slightly after this and becomes constant again around 130 deg. It is felt this indicates that the type of skin-friction distribution associated with the critical range was beginning to develop.

3 For the runs with N_{Re} = 140,000, 170,000, and 219,000, the pressure-distribution curves indicate that the second type of skin-friction distribution was present.

An example of this second type of skin-friction distribution (from ref. 5) for N_{Re} = 168,000 but with a turbulence grid in the flow before the test cylinder is plotted in Fig. 8. The experimentally determined distributions of $(N_{Nu})_\theta$ and $(p - p_0)/[(1/2)\rho_0 u_0^2]$ at N_{Re} = 170,000 are also plotted for comparative purposes. Although no turbulence grid was used in this investigation, the relatively high level of turbulence in the wind tunnel used (turbulence factor = 2.25) in some measure justifies the comparison of Fig. 8.

Assuming that the skin-friction intensity curve in Fig. 8 can be used as an indication of the skin-friction distribution which was present around the test cylinder during the runs with N_{Re} = 170,000, some interesting correlations can be noted. First of all, the location of the first minimum of the $(N_{Nu})_\theta$, or essentially h_θ, curve occurs near the minimum of the skin-friction curve. Then the location of the maximum of $(N_{Nu})_\theta$ or h_θ is approximately the same as the maximum of the skin friction. Finally, the location of the second minimum h_θ coincides approximately with the point where the skin friction drops to zero.

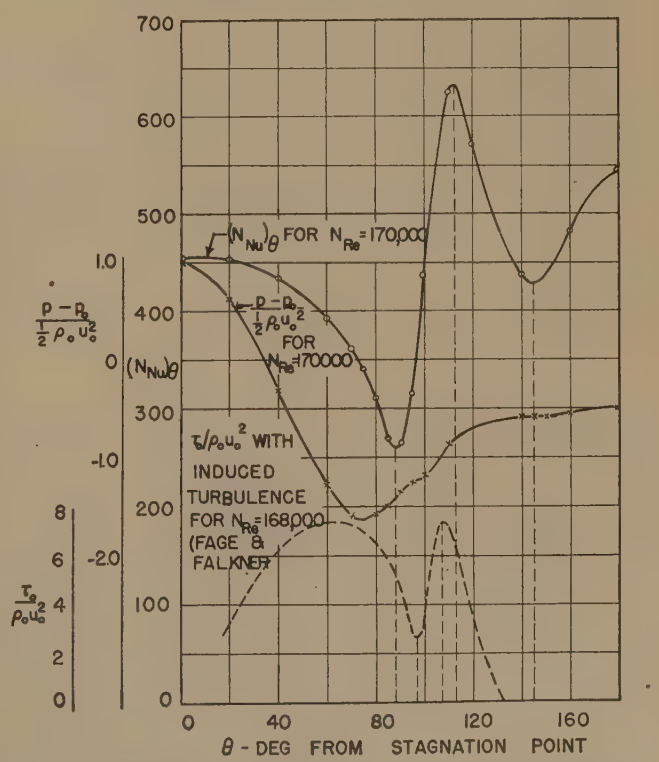


FIG. 8 DISTRIBUTION OF NUSSELT MODULUS AND PRESSURE AT REYNOLDS MODULUS = 170,000, AND DISTRIBUTION OF SKIN FRICTION AT REYNOLDS MODULUS = 168,000 WITH TURBULENCE GRID (4) OVER CIRCUMFERENCE OF A CYLINDER

The correlation between the rate of heat-transfer and skin-friction distributions shown in Fig. 8 could be expected to be only approximate. However, an examination of the pressure-distribution curves for the runs with N_{Re} = 140,000, 170,000, and 219,000 indicates that, if simultaneous skin-friction measurements could have been made, the correlation would have been very close. In Fig. 7 the first minimum of the $(N_{Nu})_\theta$ curves are noted to coincide in location with the first inflection points of the corresponding pressure curves. These inflection points, according to Fage and Falkner, mark the location of the minimum of the skin friction. The fact that the point unit heat-transfer coefficients drop to a minimum and then abruptly rise near these inflection points also substantiates their conclusion that the transition from laminar to turbulent flow in the boundary layer occurred at this same location.

Also to be noted in Fig. 7 is that the second inflection points, which occur between 140 and 150 deg in the pressure-distribution curves for these same values of N_{Re}, coincide with the locations of the second minimum h_θ values. On the basis of Fage and Falkner's findings, these second inflection points would also indicate the locations of second minimums or zero values of the skin friction.

Finally, since both the point unit heat-transfer coefficients and skin-friction intensities rise from coincident minimums on the front half of the cylinder and drop to coincident minimums on the back half, it is quite reasonable to assume the maximum values in between occur near the same point.

Effect of Different Temperature Gradients Along the Nichrome Ribbon

Although it would have been desirable, it was not possible to obtain heat-transfer data from an isothermal cylinder surface with the equipment used. It was possible, however, to cause a wide variation of the temperature distribution and temperature gradients around the cylinder circumference for any air-stream velocity by varying the power input to the ribbon.

The results of such a test are shown in Fig. 9. Increasing the power input by a factor of 4 caused the temperature gradients along the ribbon to increase from 4 to 8 times in value. The variation in the rate of heat-transfer distributions can be noted to be less than 5 per cent except at the minimum point, where it is about 8 per cent. Since experimental errors and air-flow fluctuations could have caused this much deviation, it appears that the effect of the temperature variation along the ribbon on the point unit heat-transfer coefficients was practically negligible. This point, however, is far from established by this meager amount of data, and it is felt further investigation should be conducted.

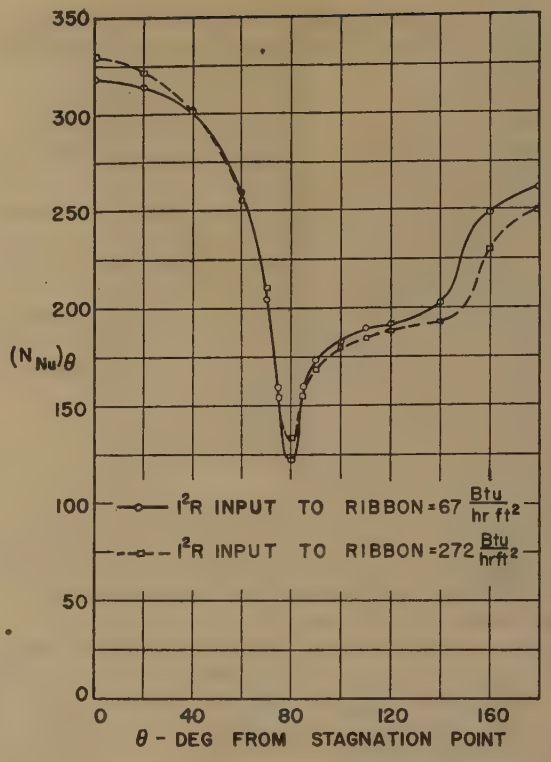


Fig. 9 Comparison of Rate of Heat-Transfer Distribution Around Cylinder at Reynolds Modulus = 101,300 With Two Different Temperature Distributions

Comparison of Results With Other Data

A survey of the measurements of the rate of heat-transfer distribution around a cylinder in an air stream made by other observers indicates that the data obtained by Schmidt and Wenner (6) were the most reliable as well as the most extensive. Therefore two experimental runs were made at values of N_{Re} (101,300 and 170,000) at which these investigators had also obtained data, so that the results could be easily compared. In Fig. 10, which shows the comparison, it can be seen that the general shape of the curves determined in this experiment is the same, but that quantitatively there is considerable deviation, particularly on the back half of the cylinder. The nonisothermal surface undoubtedly contributed to part of this, but it is felt that the widely different flow characteristics and widths of the free air streams in which the tests were conducted were also major factors. Schmidt and Wenner's data were taken in a relatively narrow jet of air with a nearly uniform velocity distribution and a minimum of turbulence, which is in direct contrast to the wind-tunnel air stream used in the present experiment.

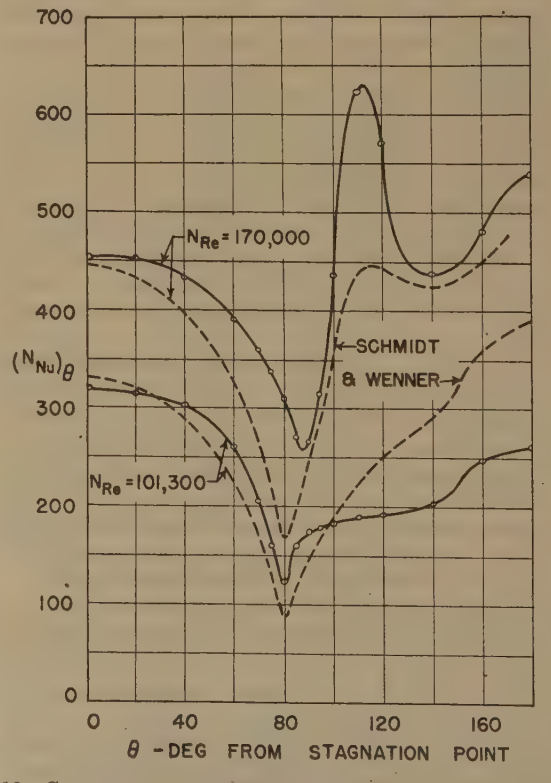


Fig. 10 Comparison of Test Data With Those of Schmidt and Wenner

Fig. 11 shows the experimentally measured values of the point unit heat-transfer coefficients at the forward stagnation point compared with the curve theoretically computed by Squire.[3] Agreement at the lower values of N_{Re} is exceptionally good; at the higher values the experimental points are about 10 per cent above the curve.

Average values of N_{Nu}, obtained from Fig. 6, are compared in Fig. 12 with Hilpert's experimentally determined average unit heat-transfer-coefficient curve for a cylinder with a constant surface temperature. Although these average values theoretically would have no significance because the cylinder surface tempera-

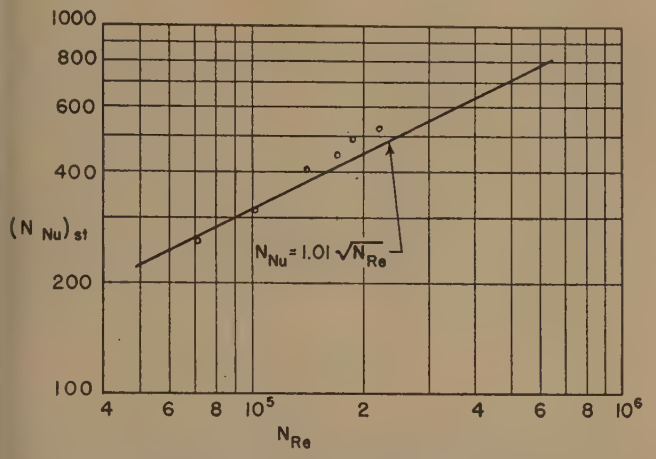


Fig. 11 Unit Heat-Transfer Coefficient at Forward Stagnation Point of Cylinder, Compared With Squire's Theoretical Solution

ture was not constant, the agreement in Fig. 12 is within 12 per cent. Again, the experimental points for the higher values of N_{Re} are high, which is due at least partly to the turbulence effect mentioned before. An important point to be noted from Fig. 12 is that the average heat-transfer coefficient around a streamlined object with an assumed constant surface temperature probably could be determined experimentally within 10 to 15 per cent accuracy from measurements of the point heat-transfer coefficient made in a manner similar to that used in this experiment.

[3] Bibliography (3), p. 631.

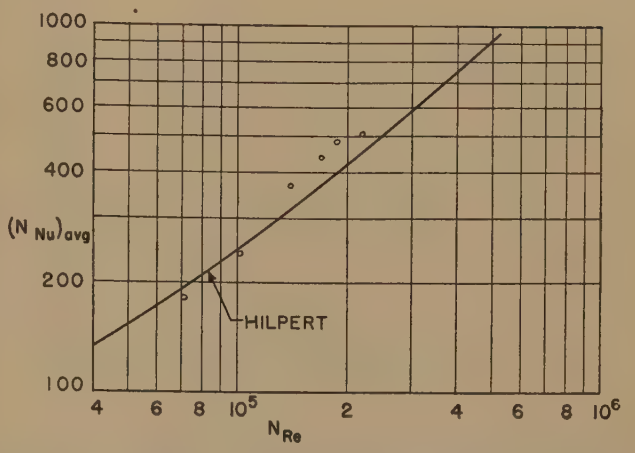


Fig. 12 Comparison of Average Heat-Transfer Coefficient for Total Cylinder Circumference With Hilpert's Experimentally Determined Curve (7)

BIBLIOGRAPHY

1 "Modern Developments in Fluid Mechanics," by the Fluid Motion Panel of the Aeronautical Research Committee, edited by S. Goldstein, Oxford University Press, London, England, 1938.

2 "An Investigation of Aircraft Heaters—NACA ARRUIII—A Simplified Method for the Calculation of the Unit Thermal Conductance Over Wings," by R. C. Martinelli, A. C. Guibert, E. H. Morrin, and L. M. K. Boelter, NACA Wartime Report W-14, 1943.

3 "Die Berechnung des Wärmeübergangs in der Laminaren Grenzschicht Umströmter Körper," by E. Eckert, *Forschungsheft* 416, VDI Verlag, Berlin, Germany, October, 1942, 26 pp.

4 "Further Experiments on the Flow Around a Cylinder," by A. Fage and V. M. Falkner, R and M No. 1369, Great Britain Aeronautical Research Committee, 1931.

5 "The Airflow Around a Circular Cylinder in the Region Where the Boundary Layer Separates From the Surface," by A. Fage and V. M. Falkner, R and M No. 1179, Great Britain Aeronautical Research Committee, 1928.

6 "Heat Transfer Over the Circumference of a Heated Cylinder in Transverse Flow," by E. Schmidt and K. Wenner, NACA TM 1050, October, 1943.

7 "Experimental Study of Heat Dissipation of Heated Wire and Pipe in an Air Current," by R. Hilpert, *Forschung auf dem Gebiete des Ingenieurwesens*, vol. 4, 1933, pp. 215–224.

8 "Heat Transmission," by W. M. McAdams, McGraw-Hill Book Company, Inc., New York, N. Y., 1942.

A Critical Review of Skin-Friction and Heat-Transfer Solutions of the Laminar Boundary Layer of a Flat Plate

By M. W. RUBESIN[1] and H. A. JOHNSON[2]

A review is made of existing literature concerned with the analytical investigation of the velocity and temperature distributions in the boundary layers of a heated(or cooled) flat plate. The plate is postulated infinitely thin and is parallel to a uniform fluid stream. The more recent solutions include the combined effects of frictional dissipation and variable fluid properties. Only the results pertaining to the transfer phenomena occurring at the plate surface are included, i.e., skin drag and over-all heat transfer; the individual temperature and velocity distributions leading to these results are omitted.

Nomenclature

The following nomenclature is used in the paper:

A = plate unit surface area for heat transfer, sq ft

C_{Dx} = local drag coefficient $= \dfrac{2\tau_x}{\rho u_0^2}$, dimensionless

c_p = specific heat of fluid at constant pressure, Btu/lb deg F

g = gravitational force per unit of mass, 32.2 ft per sec^2

h_x = local unit thermal conductance (heat-transfer coefficient), Btu/hr ft^2 deg F

J = mechanical equivalent of heat, 778 ft-lb/Btu

k = thermal conductivity of fluid, Btu/hr ft^2 (deg F/ft)

n = exponent for thermal conductivity and viscosity variation with absolute temperature, i.e.,

$$\frac{k}{k_0} = \frac{\mu}{\mu_0} = \left(\frac{T}{T_0}\right)^n, \text{ dimensionless}$$

q = heat-transfer rate, Btu per hr

r = recovery factor, dimensionless

T' = recommended fluid reference temperature, deg F abs

T_{aw} = adiabatic surface temperature, deg F abs

T_w = plate surface temperature, deg F abs

T_0 = free-stream fluid temperature, deg F abs

u_0 = free-stream fluid velocity, fps

x = distance along plate from leading edge, ft

y = distance from plate surface normal to plate, ft

$\mathbf{M}$ = Mach number, dimensionless

N_{Pr} = Prandtl modulus, $\dfrac{3600 g \mu c_p}{k}$, dimensionless

N_{Re} = Reynolds modulus $= \dfrac{u_0 x \rho}{g\mu}$, dimensionless

ρ = mass density of fluid, lb-sec/ft^4

τ_x = local drag force per unit area, psf

μ = absolute viscosity of fluid, lb-sec/ft^2

Introduction

Frictional dissipation in the boundary layers of bodies in high-speed fluid streams produces such large temperature ranges that the effect of the variation of fluid properties must be considered in analysis of skin drag and heat transfer. Existing experimental data of the characteristics of the laminar and turbulent boundary layers are of insufficient scope to shed any light on these phenomena. Only the laminar boundary layer has been approached analytically, to varying degrees of approximation, by authors in the past. A review of knowledge concerning the laminar boundary layer of a flat plate resulting from these existing analytical solutions constitutes a major portion of this paper. The remainder of the paper is devoted to analyzing these results in order to determine an inherent basis for their behavior. The results applicable to air flow are stressed.

Before presenting the particular limitations of each of the existing solutions, a formal presentation of the problem is necessary. Rigorously, the problem is as follows. The characteristics of the laminar boundary layers of temperature and velocity for an infinitely thin flat plate placed parallel to the fluid stream are found by the simultaneous solution of the following conservation of mass, momentum, and energy equations

$$\frac{\partial}{\partial x}(\rho u) + \frac{\partial}{\partial y}(\rho v) = 0 \quad \ldots\ldots\ldots\ldots [1]$$

$$\rho u \frac{\partial u}{\partial x} + \rho v \frac{\partial u}{\partial y} = \frac{\partial}{\partial y}\left(\mu \frac{\partial u}{\partial y}\right) \quad \ldots\ldots\ldots\ldots [2]$$

$$\rho u \frac{\partial}{\partial x}(c_p T) + \rho v \frac{\partial}{\partial x}(c_p T) = \frac{\partial}{\partial y}\left(k \frac{\partial T}{\partial y}\right) + \mu \left(\frac{\partial u}{\partial y}\right)^2 \quad \ldots [3]$$

with the boundary conditions

$x = 0$	$y \geq 0$	$u = u_0$	$T = T_0$
$x \geq 0$	$y = 0$	$u = 0, v = 0$	$T = T_w$ or $\dfrac{\partial T}{\partial y} = 0$
$x \geq 0$	$y \to \infty$	$u = u_0$	$T = T_0$
$x \to \infty$	$y \geq 0$	$u = 0, v = 0$	$T = T_w$ or $\dfrac{\partial T}{\partial y} = 0$

(i.e., assumption of infinitely thick boundary layers at $x \to \infty$).

It will be noticed from the manner in which the equations are written that ρ, μ, k, and c_p may vary throughout the region of integration, thus making the equations perfectly general for the case of zero pressure drop along the plate. The second term of the right member of Equation [3] is that which accounts for the dissipation of mechanical energy into thermal energy by the viscous forces prevailing in the boundary layer.

[1] Aeronautical Research Scientist, National Advisory Committee for Aeronautics, Ames Laboratory, Moffett Field, Calif.

[2] Associate Professor, Mechanical Engineering, University of California, Berkeley, Calif. Mem. ASME.

Contributed by the Heat Transfer Division and presented at the Heat Transfer and Fluid Mechanics Institute, Los Angeles, Calif., June 21, 1948. This study was made as part of the University of California Aerodynamic Heating Project as sponsored by the Aircraft Laboratory, Engineering Division, Army Air Forces, Dayton, Ohio.

From the boundary conditions, it may be observed that a certain consistency exists with respect to a variable made up of terms such as y/x^m, where m is some positive constant. When this variable approaches infinity through y approaching infinity or x approaching zero, the same conditions prevail for u and T. When the variable approaches zero through a variation of x and y, it is found that the boundary conditions are again the same. This particular point is the crux of the problem in that it enables the partial differential equations to be reduced to total differential equations which can then be integrated. A most excellent review of the mathematics of the following solutions may be found in a report by J. A. Lewis (1).[3]

The first solution employing this method for changing the partial differential equations to total differential equations was that of Blasius (2) in 1908. He determined the boundary-layer velocity distribution, and resulting drag, by integrating numerically the total differential equation resulting from the transformation of Equations [1] and [2] to a single independent variable, $y/\sqrt{x}$. By neglecting Equation [3] he, in effect, postulated a constant temperature throughout the flow boundary layer, therefore justifying his use of constant viscosity and density in Equations [1] and [2].

In 1921 Pohlhausen (3) reduced Equation [3] to a total differential equation by employing the independent variable used by Blasius. He integrated this equation employing the Blasius velocity distributions for two particular cases. First, he solved Equation [3] for the temperature distribution, neglecting frictional heating, that is, the second term of the right member, with the boundary condition that the plate has any temperature. Then he solved the case including frictional heating for the conditions where the plate is considered insulated and nonradiating. In both of the foregoing solutions, by use of the Blasius velocity distribution, Pohlhausen, of necessity, considered the fluid properties constant throughout the flow and thermal boundary layers. The more general case, that with frictional dissipation occurring simultaneously with heat transfer but still based on the Blasius velocity distribution, was solved analytically by Eckert and Drewitz (4) in 1940.

The neglect of the variation of fluid properties in the boundary layers by the foregoing solutions leaves doubt of their applicability to high-velocity cases because of the large temperature ranges which prevail whenever frictional dissipation is important. The solutions which will be outlined attempt to rectify this matter. It should be noted that in every case the method of converting the partial to total differential equations is similar to that used by Blasius.

In 1935 Busemann (5) set forth the procedure for integrating Equations [1] to [3] for the conditions that $N_{Pr} = 1$ and that the viscosity and thermal conductivity vary with the square root of the absolute temperature as predicted from kinetic theory. The specific heat is considered invariant. Only one particular case is integrated, however, this being that of an insulated plate and gas flow at $\mathbf{M} = 8.8$.

An analysis was performed in 1939 by von Kármán and Tsien (6) of Equations [1] to [3] for the conditions that the viscosity and thermal conductivity vary with the 0.765 power of the absolute temperature (experimentally valid for diatomic gases) and $N_{Pr} = 1$, that is, the specific heat is constant. The following boundary conditions were investigated:

1 No heat transferred to or from the wall.
2 Wall temperature equal to $^1/_4$ free-stream temperature.

Solutions of temperature and velocity distribution are included for a Mach number range of 0 to 10.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

The analyses of Busemann and of von Kármán and Tsien are severely restricted in that N_{Pr} was maintained equal to 1. Since the N_{Pr} of air is approximately 0.76, there was some question as to the validity of applying the results of these analyses to air flow. Cognizance of this limitation led various authors to attempt to remedy it by performing analyses where the Prandtl modulus was made approximately equal to that of air. These attempts were entirely independent, but it was found, curiously enough, that the earliest analysis was the most straightforward and complete.

The analysis last mentioned was performed by Crocco (7) in 1941. Aid in the numerical integration of his reduced equations was provided by Conforto (8). The analysis included simultaneous solution of the boundary-layer equations for the following conditions:

(*a*) $N_{Pr} = 0.725$
(*b*) Exponent for variation of viscosity and thermal conductivity, $n = 1.25;\ 1.00;\ 0.75;\ 0.5$

The range of Mach numbers analyzed is from 0 to 5, and the surface-temperature range considered is from 0.25 to about twice the free-stream temperature. It should be noted that the ranges of the variables were not extended because the authors felt that the postulate of a constant exponent for the viscosity and thermal-conductivity variation with temperature would not hold throughout the entire boundary layer at higher values of the variables, $\mathbf{M}$ and T_w/T_0

In 1942 Hantzsche and Wendt (9) summarized their work with the variable-fluid-property laminar boundary layer. To facilitate integration of their reduced equations (same as used by Crocco), they limited most of their solutions to the cases where $N_{Pr} = 1$ and the viscosity exponent was 0.5 and 0.8, or where the viscosity exponent was 1.0 and N_{Pr} was variable. Only one case, where $N_{Pr} = 0.7$ and $n = 0.8$, was integrated numerically for the conditions of an insulated plate at an air flow of $\mathbf{M} = 5.4$. The latter analysis seemed to indicate that the velocity and temperature distributions were well satisfied for this one case by the more easily integrated case of $N_{Pr} = 1$. An important point to note was that the heat-transfer and drag coefficients were independent of the temperature ratio T_w/T_0, for conditions where $n = 1.0$ and N_{Pr} had any value. Thus the effects of the wall temperature and Mach number, believed to be important, are masked by the case where $n = 1$

Probably the most elegant of existing solutions, unfortunately restricted to an insulated plate, is that of Brainerd and Emmons (10) in 1942. These authors employed a differential analyzer to integrate their reduced boundary-layer equations and consequently, may have more accurate answers than those of the analyses presented previously. Their most applicable results for air are for $N_{Pr} = 0.733$ and viscosity exponent, n equal to 0.768. The range of $\mathbf{M}$ for their computations are from 0 to 3.16.

Discussion of Results

This discussion of the results of the analyses presented in the foregoing will be limited to those cases which approximate air flow. To systematize the discussion, it is divided into the following main categories:

1 General results common to most of the analyses.
2 Results pertinent to an insulated flat plate, i.e., recovery factor and coefficient of drag.
3 Results pertinent to a flat plate with heat transfer through its surface, i.e., the relationship between the Nusselt modulus and the coefficient of drag, and the influence of the various parameters of the solution on the latter.
4 Reference temperature for evaluating the fluid properties so that the effect of the various parameters is eliminated.

General. A most important characteristic common to the foregoing analyses which treat with heat transfer is that the effect of the frictional heating is to modify the temperature potential in the Newton equation

$$q/A = h(T_w - T_0) \quad [4]$$

while leaving the heat-transfer coefficient at a value approximately that of low-speed fluid flow. This was shown analytically in algebraic form by Eckert and Drewitz, and was substantiated numerically by the other heat-transfer analyses.

The temperature potential is modified to the extent that the free-stream temperature (T_0) is replaced by the temperature that the surface would attain due to frictional heating while insulated from its surroundings. This latter temperature is henceforth called the adiabatic surface temperature or the effective ambient-air temperature.

Insulated Plate. The adiabatic surface temperature, defined previously, can be expressed conveniently by the equation

$$T_{aw} = T_0 + r \frac{u_0^2}{2gJc_p} \quad [5]$$

where r is the recovery factor. An important point, characteristic of all the solutions, is that the recovery factor is independent of the Mach and the Reynolds modulus, depending solely on the Prandtl modulus. Pohlhausen and Eckert and Drewitz found that the recovery factor could be approximated by $(N_{Pr})^{1/2}$ for values of N_{Pr} between 0.5 and 1.2. The work of Busemann and von Kármán and Tsien could not contribute in the determination of the exponent for the modulus as N_{Pr} was maintained equal to 1 in their analyses. For a linear variation of viscosity and thermal conductivity ($n = 1$), Hantzsche and Wendt also indicated that $(N_{Pr})^{1/2}$ sufficed as a recovery factor. The numerical solution of Crocco-Conforto showed for $N_{Pr} = 0.725$ that the recovery factor was 0.85 which is approximately equal to $N_{Pr}^{1/2}$ (to 0.1 per cent). The Brainerd and Emmons solution, however, resulted in a value of recovery factor about 1.2 per cent less than $N_{Pr}^{1/2}$ One must not have too much credence in the latter result, however, because the value obtained for $N_{Pr} = 1$ was also low by 0.7 per cent.

The quantity $C_{Dx}\sqrt{N_{Re}}$ for the various solutions is plotted as a function of Mach number in Fig. 1. This quantity is used because it is constant and equal to 0.664 (Blasius) for constant-property fluid flow, and therefore provides a means for representing the variations due to variable properties. A comparison of the von Kármán and Tsien, and the Brainerd and Emmons solutions indicates that the effect of the variation of the Prandtl modulus from 0.733 to 1 is negligible. The effect of n, the exponent of temperature in the viscosity and thermal-conductivity relationships, is much greater, as indicated by comparison of the Hantzsche and Wendt, and Busemann solutions with that of von Kármán and Tsien. In fact, Hantzsche and Wendt indicated that when $n = 1$ the result obtained is identical to that of the constant-property case. The results of Crocco and Conforto are uniformly lower than those of von Kármán and Tsien; however, this is not ascribed to the small differences in the values of the parameters as they have no effect at the $\mathbf{M} = 0$ point which is also low. In general, the quantity $C_{Dx}\sqrt{N_{Re}}$ decreases with increasing Mach number, the variation from the Blasius solution being about —14.5 per cent at Mach = 5 and —24.7 per cent at Mach = 10. It should be noted that the temperature at which the properties of the coefficient of drag and the Reynolds modulus are evaluated is the free-stream temperature. Later a discussion of a more fundamental temperature at which these properties may be evaluated will be made

Plate With Heat Transfer. The relationship between the heat-transfer coefficient and the drag coefficient for a plate from which heat is being transferred is derived in the Appendix. The expression is

$$\frac{h_x x}{k} = \frac{C_{Dx}\sqrt{N_{Re}}}{2} N_{Pr}^{1/3} N_{Re}^{1/2} \quad [6]$$

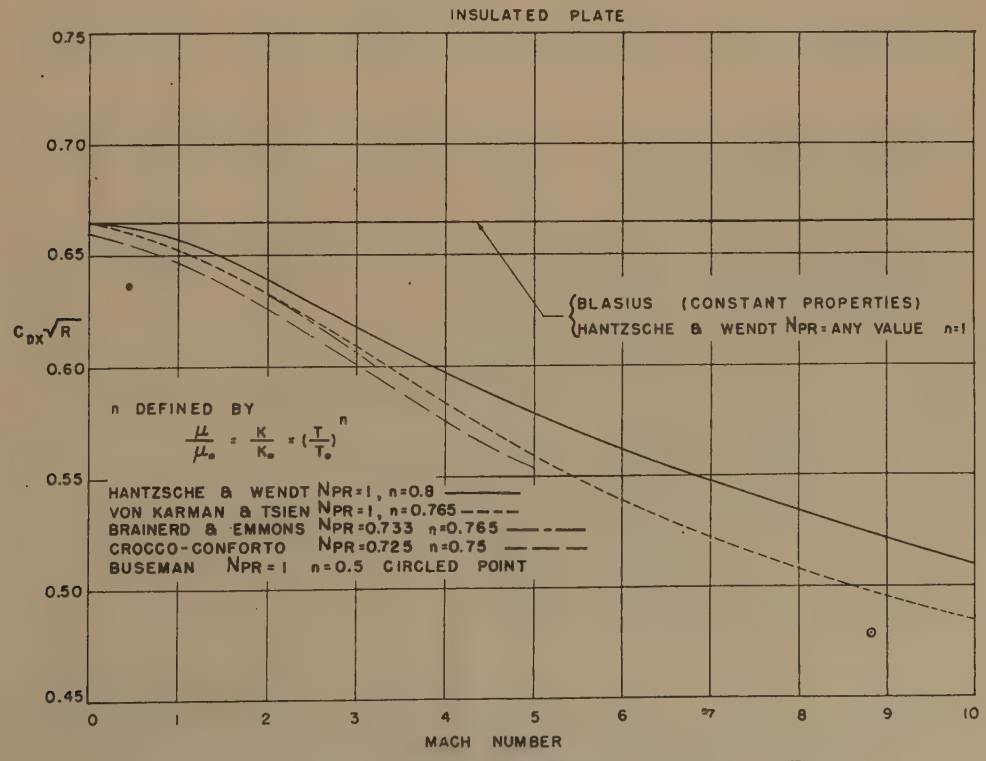


FIG. 1 VARIATION OF FLAT-PLATE DRAG COEFFICIENT; INSULATED PLATE

The simplicity of this expression results from defining the coefficient of drag so that it is based upon a gas density evaluated at the same temperature as the rest of the fluid properties not necessarily the free-stream temperature as in the usual definition. Thus, once $C_{Dx}\sqrt{N_{Re}}$ is known, the heat-transfer coefficient also can be evaluated. The heat-transfer coefficient determined from this expression must be used with the modified temperature potential explained in the previous section. Since the fluid properties of the terms in the various parameters of Equation [6] are evaluated at the same reference temperature, when Fig. 2 is used this temperature must be that of the free stream. The latter is arbitrary for the Prandtl modulus because no definite reference temperature for this modulus can be deduced from the solutions as they all postulate N_{Pr} constant throughout the thermal boundary layer.

Curves of the relationship $C_{Dx}\sqrt{N_{Re}}$, based on the free-stream temperature versus Mach number, are presented in Fig. 2 for a plate from which heat is being transferred. The parameter of each curve is the ratio of wall temperature to free-stream temperature. The results indicated are chiefly those of Crocco and Conforto for the conditions that $N_{Pr} = 0.725$ and $n = 0.75$. One curve that was computed by von Kármán and Tsien for $N_{Pr} = 1$ and $n = 0.765$ at $T_w/T_0 = 1/4$ is also included for comparison. A comparison of the two solutions indicates that the effect on $C_{Dx}\sqrt{N_{Re}}$ of the variation of the Prandtl modulus and the small difference of n is negligible. Again, the results of Crocco-Conforto are uniformly low as indicated by the divergence of their solution at $\mathbf{M} = 0$ and $T_w/T_0 = 1$ from the Blasius solution. In general, the quantity $C_{Dx}\sqrt{R}$ decreases with increasing Mach number and also with increasing surface temperature. It should be noted again from Hantzsche and Wendt that if $n = 1$, $C_{Dx}\sqrt{N_{Re}} = 0.664$ and is independent of Mach number or the surface temperature.

Reference Temperature for Fluid Properties. Since the definitions of drag coefficient and Reynolds modulus are arbitrary in terms of the temperature at which to evaluate the fluid properties, there may be some temperature in the boundary layer at which these properties can be evaluated so as to eliminate the effect of small variations of the variables N_{Pr} and n and also make $C_{Dx}\sqrt{N_{Re}}$ independent of Mach number and surface temperature. The advantage of this is that drag and heat transfer could be expressed mathematically in such form as to eliminate the need of using curves necessary for accounting for variations in N_{Pr}, which has little influence, and n, $\mathbf{M}$, and T_w, which have large influence.

A temperature expression which satisfied the foregoing requirements was found from the results of Crocco-Conforto indicated in Fig. 2. This was performed by solving graphically for the temperature required to eliminate the effects of the parameters and then expressing the mean values of the solutions by a mathematical expression, namely

$$\frac{T'}{T_0} = 1 + 0.032\,\mathbf{M}^2 + 0.58\left(\frac{T_w}{T_0} - 1\right) \quad [7]$$

As the insulated plate is just a special case of the plate with general surface temperature, Equation [7] should apply for the insulated plate when T_w/T_0 is replaced by T_{aw}/T_0 (see Equation [5]). When this is done, Equation [7] reduces to an expression equal to

$$\frac{T'}{T_0} = 1/4(1 + 3\,T_{aw}/T_0) \quad [8]$$

Equation [8], surprisingly, is the same as that obtained when the surface temperature is weighted twice in determining an arithmetic average of the wall temperature and the free-stream temperature.

It can be shown that the expression for drag coefficient and Reynolds modulus based upon a temperature T' can be evaluated from that based upon the free-stream temperature through use of the expression

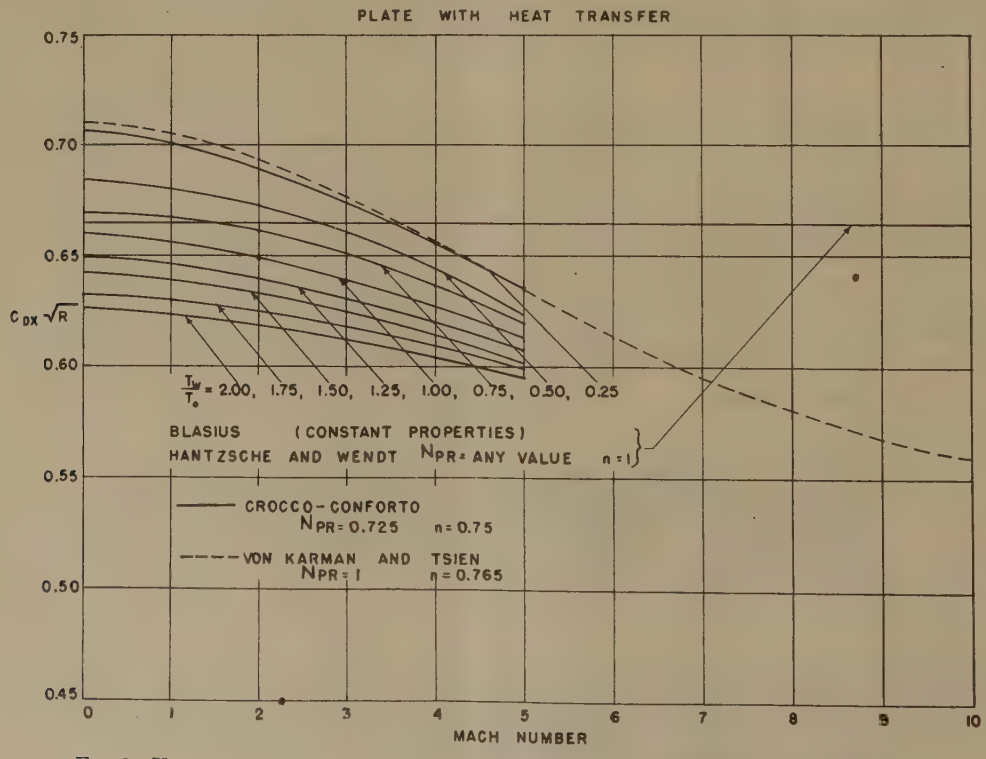


FIG. 2 VARIATION OF FLAT-PLATE DRAG COEFFICIENT; PLATE WITH HEAT TRANSFER

$$C'_{Dx}\sqrt{N'_{Re}} = (T'/T_0)^{\frac{1-n}{2}} C_{Dx}\sqrt{N_{Re}} \quad \ldots\ldots\ldots [9]$$

By employing the relations, Equations [7], [8], [9], Figs. 1 and 2 are replotted in Figs. 3 and 4. It is observed that the desired result of eliminating the effect of the various parameters is attained approximately for both the insulated and heated (or cooled) plate. The deviations from the constant-property case by some of the solutions can be explained in that the relations, Equation [7] and [8] were derived only from the Crocco-Conforto solutions and arbitrarily applied to the others. In general, it can

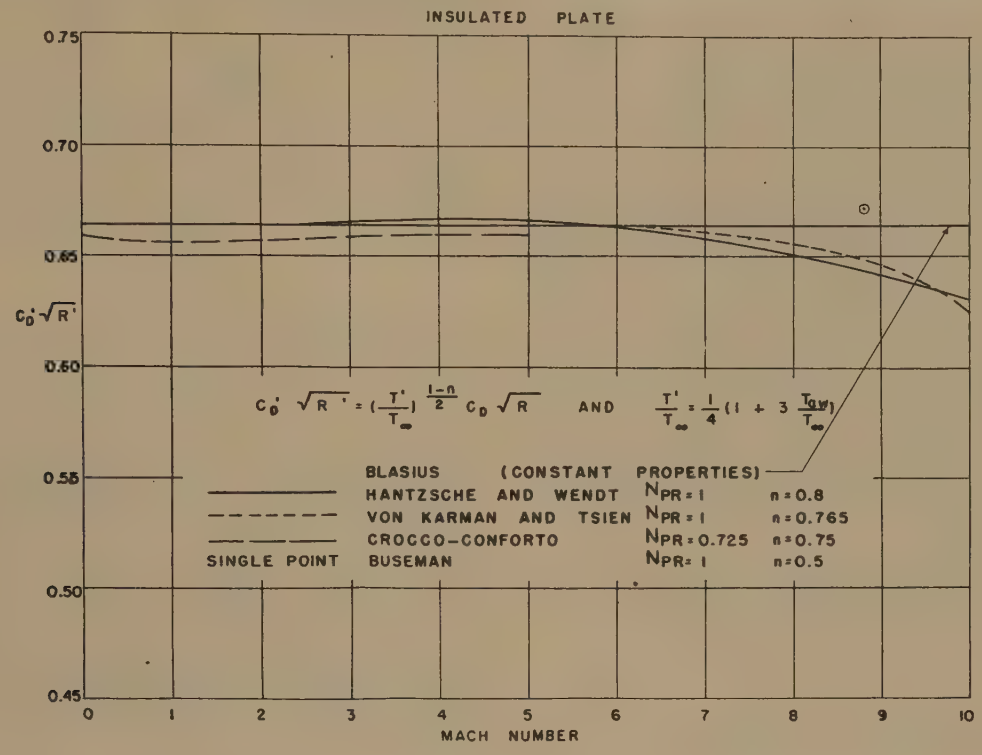


FIG. 3 VARIATION OF FLAT-PLATE DRAG COEFFICIENT; INSULATED PLATE

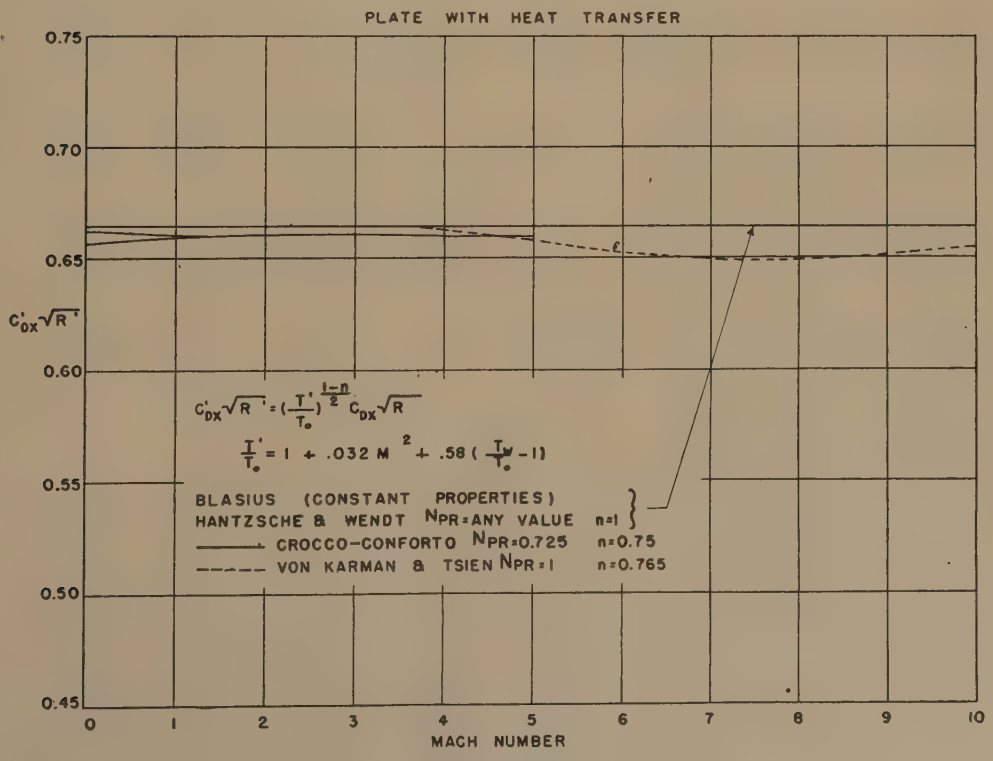


FIG. 4 VARIATION OF FLAT-PLATE DRAG COEFFICIENT; PLATE WITH HEAT TRANSFER

be concluded that Equations [7] and [8] satisfy the conditions desired exactly to $\mathbf{M} = 5$ and approximately to $\mathbf{M} = 10$.

BIBLIOGRAPHY

1 "Monograph V, Boundary Layer in Compressible Fluid," by J. A. Lewis, AMC-T2 Technical Report F-TR-1179-ND, February, 1948.

2 "Grenzschichten in Flüssigkeiten mit kleiner Reibung," by H. Blasius, *Zeitschrift für Mathematik und Physik*, vol. 56, 1908, p. 1 (see p. 84 Durand's Aerodynamic Theory, vol. 3).

3 "Der Wärmeaustausch zwischen festen Körpern und Flüssigkeiten mit kleiner Reibung und Kleiner Wärmeleitung," by E. Pohlhausen, *Zeitschrift für angewandte Mathematik und Physik*, vol. 1, 1921, pp. 115–121.

4 "The Heat Transfer to a Plate in Flow at High Speed," by E. Eckert and O. Drewitz, NACA Technical Memo no. 1045, May, 1943.

5 "Laminar Flow of a Gas Along a Plate," by A. Busemann, *Zeitschrift für angewandte Mathematik und Physik*, vol. 15, 1935, p. 23.

6 "Boundary Layer in Compressible Fluids," by T. von Kármán and H. S. Tsien, *Journal of Aeronautical Sciences*, vol. 5, 1938, pp. 227–232.

7 "Boundary Layer of Gases Along a Flat Plate," by L. Crocco, Seminario Matenatico Rendiconti, University of Roma, Rome, Italy, vol. 2, 1941, p. 138.

8 "The Integration of a System of Equations Relating to the Theory of the Boundary Layer in Gases," by F. Conforto, Seminario Matenatico Rendiconti, University of Roma, Rome, Italy, vol. 2, 1941.

9 "The Laminar Boundary Layer of the Flat Plate With and Without Heat Transfer Considering Compressibility," by W. Hantzsche and H. Wendt, Jahrbuch 1942 der deutschen Luftfahrtforschung, part 1, pp. 40–50.

10 "Effect of Variable Viscosity on Boundary Layers, With a Discussion of Drag Measurements," by J. G. Brainerd and H. W. Emmons, *Journal of Applied Mechanics*, Trans. ASME, vol. 64, 1942, p. A-1.

Appendix

The expression for the relationship between the coefficient of drag and heat-transfer coefficient is derived as follows

By definition

$$\frac{q_x}{A} = h_x(T_w - T_{aw}) = -k_w \left.\frac{\partial T}{\partial y}\right|_w \quad \ldots\ldots [10]$$

and

$$\tau_x = \mu_w \left.\frac{\partial u}{\partial y}\right|_w = \frac{c_{Dx}}{2} \rho u_0^2 \quad \ldots\ldots [11]$$

Note that the ρ is left at a temperature T and not the free-stream temperature T_0 as is usually done.

From Equation [10]

$$h_x = \frac{-k_w \left.\frac{\partial T}{\partial y}\right|_w}{(T_w - T_{aw})} = -\frac{k_w \left.\frac{\partial T}{\partial u}\right|_w \left.\frac{\partial u}{\partial y}\right|_w}{(T_w - T_{aw})} \quad \ldots\ldots [12]$$

Introducing Equation [11] into Equation [12]

$$h_x = \frac{-k_w \left.\frac{\partial T}{\partial u}\right|_w \frac{c_D}{2} \rho u_0^2 \frac{1}{\mu_w}}{(T_w - T_{aw})} \quad \ldots\ldots [13]$$

Multiplying Equation [13] by x/k_w results in

$$\frac{h_x x}{k_w} = \frac{-\left.\frac{\partial T}{\partial u}\right|_w \frac{c_D}{2} \rho u_0 x \frac{u_0}{\mu_w}}{(T_w - T_{aw})} \quad \ldots\ldots [14]$$

In order that Nusselt modulus be based upon a reference temperature T, multiply the left member of Equation [14] by k_w/k and the right member by μ_w/μ. Note that $k_w/k = \mu_w/\mu$.

$$\frac{h_x x}{k} = -\left\{\left.\frac{\partial T}{\partial u}\right|_w \frac{u_0}{(T_w - T_{aw})}\right\} \frac{c_{Dx}}{2} \frac{\rho u_0 x}{\mu} \quad \ldots\ldots [15]$$

The bracketed term can be shown from the analytical result of Eckert and Drewitz (3) to be approximately equal to $N_{Pr}^{1/3}$. This was further substantiated numerically by the other analyses. Therefore

$$\frac{h_x x}{k} = \left(\frac{c_{Dx}\sqrt{N_{Re}}}{2}\right) N_{Pr}^{1/3} N_{Re}^{1/2}$$

where properties are evaluated at temperature T.

Torsional Viscous-Friction Dampers

By J. C. GEORGIAN,[1] MILWAUKEE, WIS.

This paper describes the large torsional viscous-friction or "Lanchester" dampers manufactured by the author's company for use on Diesel engines. Some properties of the new silicone fluids, that are necessary for damper design, are given. The theory of the torsional viscous-friction damper as applied to multimass systems is developed. Constructional details of the damper and methods of manufacture are discussed. Some test results are given and compared with calculations.

Introduction

SINCE Lanchester's patent (1)[2] on viscous-friction dampers, there have been innumerable variations of his patent for both viscous- and dry-friction dampers. Owing to the fact that in the early days there was no suitable viscous fluid, most dampers were of the spring-loaded dry-friction type. Most of these devices suffer from the defect that their usefulness is dependent upon close adjustment of springs, which adjustment rarely is at the optimum point after a period of operation.

Around 1930 Dashefsky (2) developed a viscous-friction damper similar in construction to the damper described in this paper. For his viscous medium he used a furniture glue. This damper was quite successful for the use for which it was intended.

With the wartime development of the silicone oils (3, 4, 5), there became available a new fluid with desirable properties necessary for the successful operation of a viscous-friction damper. These properties and characteristics will be discussed in the next section.

Silicone Fluid

The desirable properties for the fluid for a viscous-friction damper are that the fluid should be stable and noncorrosive at fairly high temperature and should have a relatively flat viscosity-temperature curve over a wide range and, further, the viscosity should not change appreciably with high rates of shear.

Fig. 1 shows the viscosity - temperature curve for silicone fluids produced by General Electric and Dow Corning Corporation. Fig. 2 shows the variations of viscosity with shear rate for General Electric and Dow Corning fluid. Table 1 is a tabulation of the properties of the Dow Corning 200–30,000 and General Electric 9996–26,000 silicone oil necessary for damper design.

TABLE 1 PROPERTIES OF SILICONE OIL

	Dow Corning	General Electric
Specific gravity	0.974	0.969
Coefficient of volume expansion	0.96×10^{-3} deg C	0.91×10^{-3} deg C

An undesirable property of silicone oil is its lack of oiliness under thin-film conditions with certain combinations of materials. Reference (4) gives lists of good, indifferent, and poor combinations of metals using silicone oil as a lubricant. Especially poor is the combination of steel on steel or steel on cast iron. This will be discussed more in detail under manufacture of the damper.

Description of Damper

From Figs. 3, 4, and 9, it may be seen that the torsional viscous-friction damper consists of a flywheel hermetically sealed within a housing with clearances between the flywheel and the housing. This clearance space between the flywheel and housing is filled with the silicone oil. During normal operation, the flywheel is dragged along at the same revolutions per minute as the engine. At a critical speed, however, the crankshaft vibrates and as the housing is securely attached to the crankshaft, it takes up the vibratory motion of the crankshaft, and as the flywheel tends to rotate uniformly, there is a relative motion between the flywheel and housing setting up opposing shear forces in the silicone fluid which tend to damp the vibration. As can be seen, the damper is the utmost in simplicity and reliability, it requires no maintenance, and is foolproof for the life of the engine.

Designing the Damper

The damper size depends on the size of the engine, and the size of the "critical" that must be damped out. The following equation has been developed in the Appendix

$$\tau_{damp} = \frac{\tau_{max}}{1 + \dfrac{125{,}000\, I_d}{\tau_{max}\, \Sigma I\theta^2}} \qquad [25]$$

where

τ_{damp} = damped vibratory stress in crankshaft
τ_{max} = vibratory stress in crankshaft with engine damping only
I_d = polar moment of inertia of damper mass
$\Sigma I\theta^2$ = Effective moment of inertia of system referred to free end of crankshaft. This is the arithmetic sum of the product of the individual masses and the square of its amplitude from the normal elastic curve or Holzer tabulation. (This includes the damper housing plus one half of the damper mass.)

Equation [25] gives the stress in the crankshaft with the damper operating, in terms of the stress in the crankshaft without the damper, and the ratio of the damper inertia to the equivalent inertia of the system. The required damper inertia must be selected so as to reduce the maximum stresses in a given system to some allowable value. The damper inertia usually will be between 10 and 25 per cent of the equivalent inertia of the system, depending upon the severity of the "critical" to be damped. The value of τ_{damp} should not exceed 4000 psi as the upper limit.

The use of Equation [25] will be illustrated by an example: A 7-cylinder four-cycle supercharged $16^1/_2$-in. $\times$ $24^1/_2$-in. Diesel engine driving a pipe-line pump has a severe 7th-order two-node vibration within the operating range, that we wish to damp out with a viscous-friction damper. The calculated maximum stress $\tau_{max} = 8290$ psi, and the effective moment of inertia is $\Sigma I\theta^2 = 2956$ lb-in-sec^2. As the 7th-order critical in a 7-cylinder engine is a severe major critical, we shall select I_d to be approximately 25 per cent of $\Sigma I\theta^2$, or $I_d = 750$ lb-in-sec^2. Substituting these values in Equation [25] we obtain

---

[1] Torsional Analyst, Nordberg Manufacturing Company. Jun. ASME.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Oil and Gas Power Division and presented at the Annual Meeting, New York, N. Y., November 29–December 3, 1948, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—A-67.

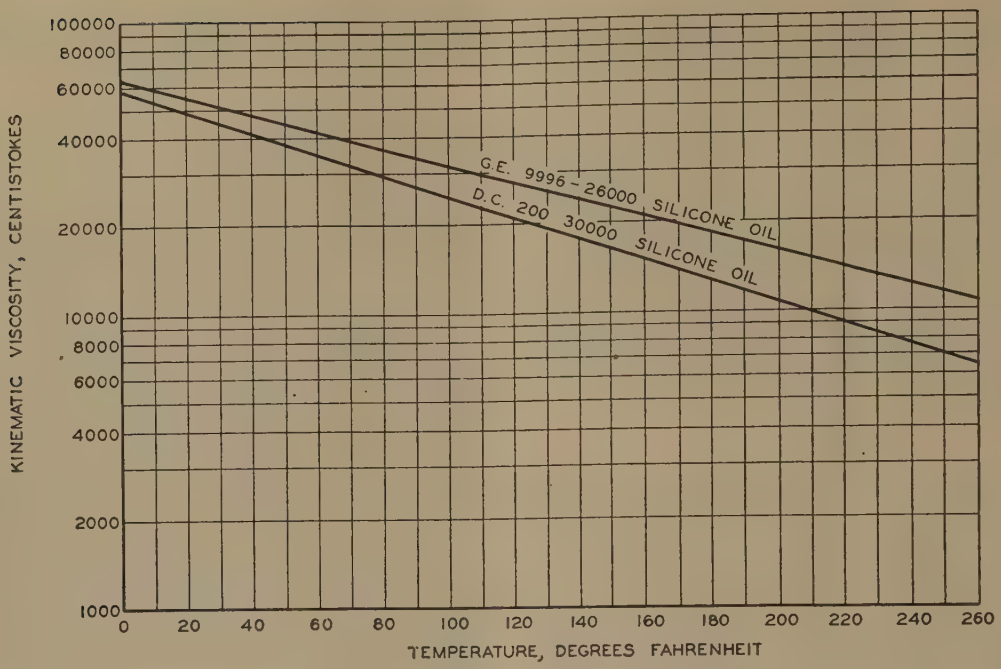


FIG. 1 VISCOSITY-TEMPERATURE CURVES OF SILICONE OIL

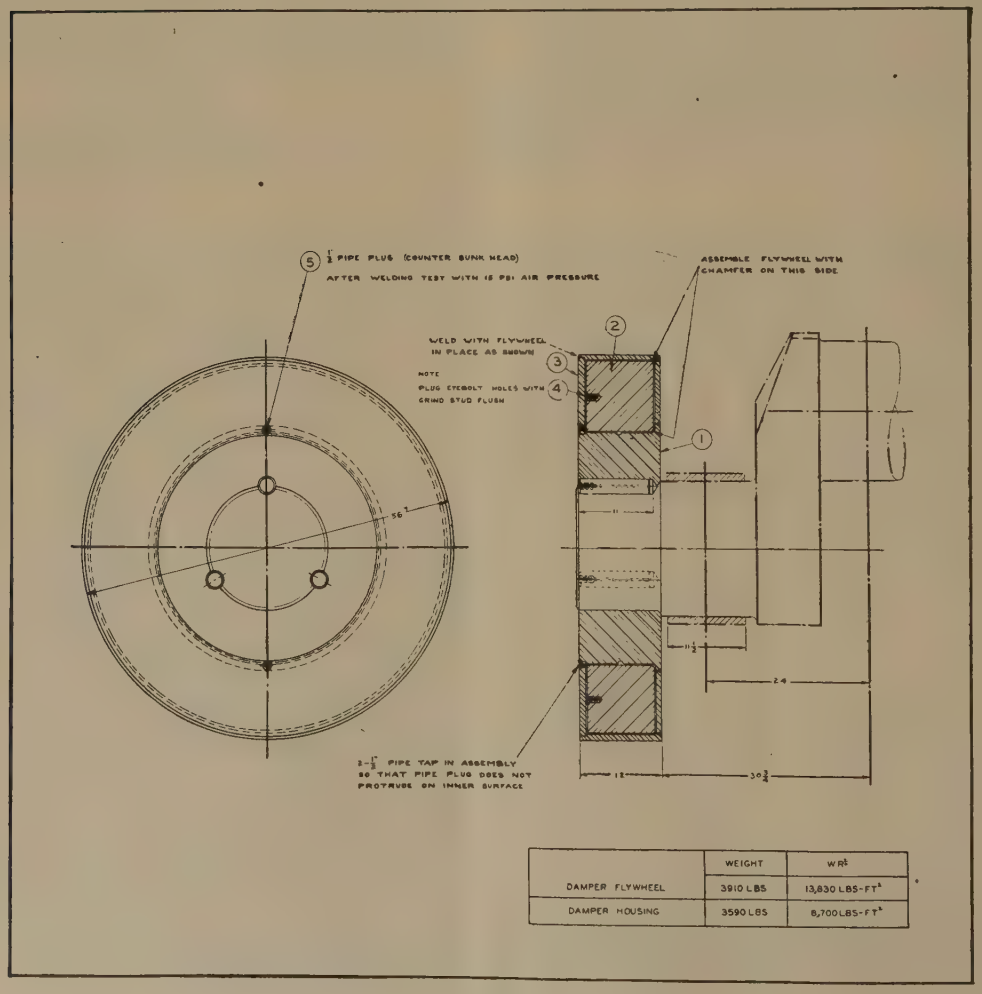


FIG. 3 ASSEMBLY OF 56-IN-DIAM DAMPER

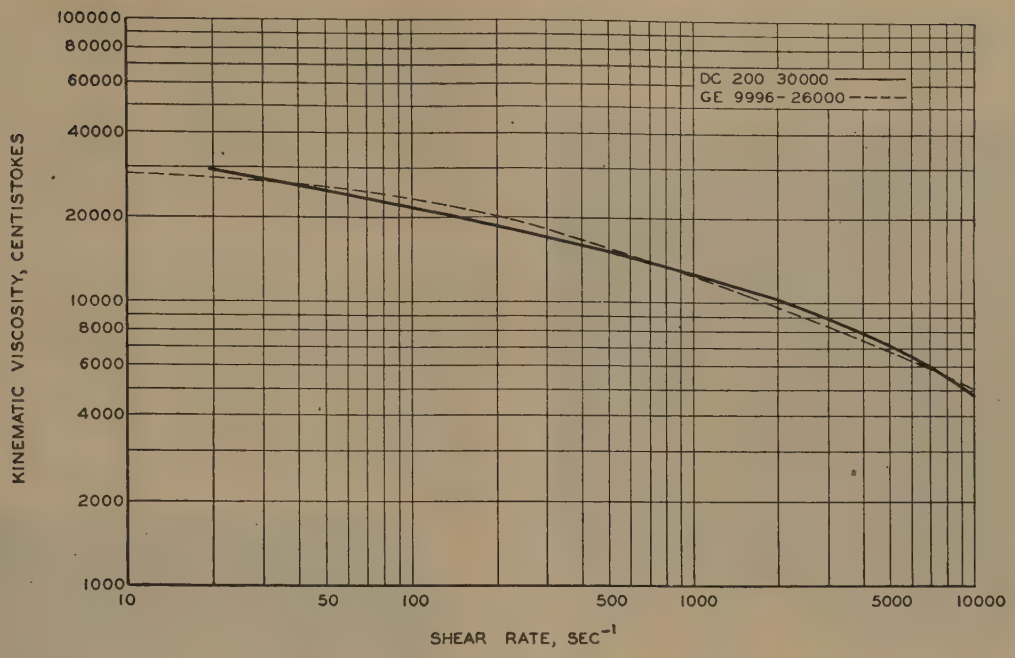


FIG. 2 VISCOSITY-SHEAR RATE CURVES OF SILICONE OIL

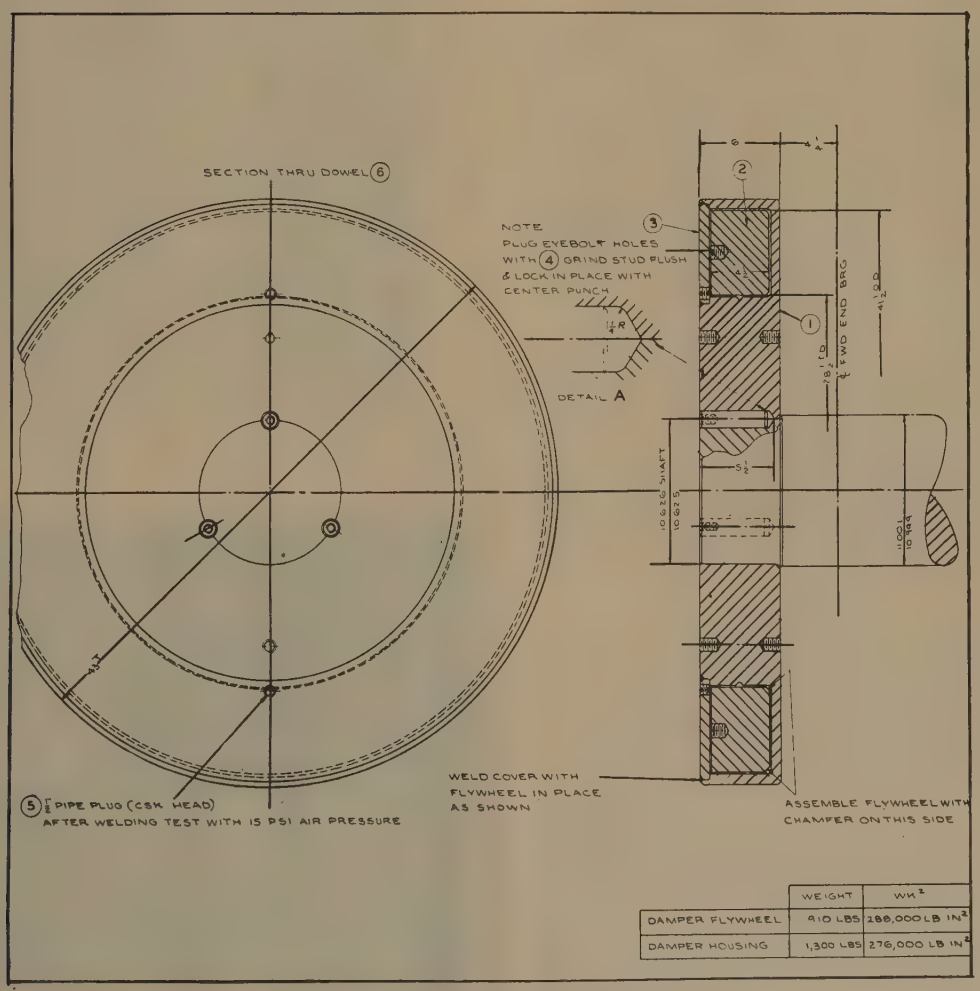


FIG. 4 ASSEMBLY OF 43-IN-DIAM DAMPER

$$\tau_{damp} = \frac{\tau_{max}}{1 + \frac{125{,}000\, I_d}{\tau_{max}\, \Sigma I\theta^2}} = \frac{8290}{1 + \frac{125{,}000 \times 750}{8290 \times 2956}} = 1720 \text{ psi}$$

As this stress is well below allowable limits, the damper inertia selected will be adequate for this engine.

After the selection of the size of the damper inertia, it is necessary to determine its proportions. The proportions may be determined from Equation [36] developed in the Appendix

$$\mu = \frac{\frac{\gamma h \omega R_o}{4g}}{\frac{1 + \left(\frac{R_i}{R_o}\right)^3}{1 - \left(\frac{R_i}{R_o}\right)^4} + \frac{R_o}{2b}} \qquad [36]$$

where

μ = absolute viscosity
h = oil-film thickness
ω = natural circular frequency of vibratory system
g = acceleration of gravity
R_o = outer radius of damper mass
R_i = inner radius of damper mass
b = width of damper mass
γ = density of damper mass

The equation gives a relation between the viscosity of the fluid, the thickness of the fluid, and the dimensions of the damper mass. It will be noted that in order to obtain the maximum thickness of the fluid, it is necessary for the damper to be as narrow as possible, i.e., for a given inertia, the damper mass should have a large diameter with small width. However, this usually will have to be a compromise, as a large-diameter damper at the forward end of an engine may interfere with piping, etc. With the fluids presently available, the thickness h for the fluid thickness will be of the order of 0.010 in. The working viscosity of the fluid in the damper should be of the order of 10,000 centistokes. This value of viscosity will be required for the usual working temperatures and shear rates in the normal damper.

Continuing with the foregoing example, we shall assume an oil-film thickness of 0.010 in., and after some estimating, damper-mass dimensions were selected as follows: Outer diameter $41^1/_2$ in., inner diameter $28^1/_2$ in., and width $4^1/_2$ in. With these dimensions, the polar moment of inertia of the damper mass calculates to be 748 lb-in-sec^2, which is close enough to 750 for practical purposes. The circular natural frequency of the system with the damper attached was calculated to be 163 radians per sec. Substituting these values in Equation [36], we have

$$\mu = \frac{\frac{\gamma h \omega R_o}{4g}}{\frac{1 + \left(\frac{R_i}{R_o}\right)^3}{1 - \left(\frac{R_i}{R_o}\right)^4} + \frac{R_o}{2b}} = \frac{\frac{0.283 \times 0.010 \times 163 \times 20.75}{4 \times 386}}{\frac{1 + \left(\frac{14.25}{20.75}\right)^3}{1 - \left(\frac{14.25}{20.75}\right)^4} + \frac{20.75}{2 \times 4.5}}$$

$$= 0.0015 \text{ reyns}$$

or converting to centipoises, by using the conversion factor of 1 reyn = 6.9×10^6 centipoises, we have $0.0015 \times 6.9 \times 10^6 = 10{,}350$ centipoises, and, dividing by the specific gravity, we get the kinematic viscosity of

$$\frac{10{,}350}{0.97} = 10{,}660 \text{ centistokes}$$

From Fig. 1 for 135 F, the usual operating temperature of the lubricating oil in a Diesel engine, the kinematic viscosity of silicone oil is around 20,000 centistokes. So the foregoing value of 10,660 centistokes is on the safe side.

The shear rate (Equation [38] in the Appendix) of the fluid in this damper when operating at the 7th-order critical is

$$\text{Shear rate} = \frac{\omega \theta_1 R_o}{\sqrt{2}\, h} = \frac{163 \times 0.268 \times 20.75}{\sqrt{2} \times 57.3 \times 0.010} = 1120 \text{ sec}^{-1}$$

Where $\theta_1 = 0.268$ deg is the vibratory amplitude at the damper housing, determined from the ratio of $\tau_{damp} = 1720$ psi to τ, where τ is the stress in the crankshaft in pounds per square inch per degree amplitude of the damper housing, and is calculated from the Holzer tabulation.

Entering Fig. 2 with this value of shear rate, we find the kinematic viscosity of the silicone oil is about 13,000 centistokes, so the design is still on the safe side.

Fig. 4 is the assembly drawing of the damper example just given.

Manufacture of the Damper

In order for the torsional viscous-friction damper to operate with the closely required clearances, it is necessary that the dimensions of the parts be controlled accurately, and the housing be of rigid construction. From Fig. 3 it can be seen that the damper is a welded construction of forgings and steel plates. The outer rim of the housing is seam-welded to form the outer ring. The back plate is then welded to this ring and the forged hub to form the housing. This welding is done on an automatic welding machine. The welded housing is annealed and finish-machined, and the inner surface polished.

The inertia flywheel, Fig. 5, is a mild-steel forging finish-machined and polished to close tolerances. For "criticals" of small magnitude, the shear rate will be relatively low so steel on steel using silicone oil as a lubricant will operate satisfactorily. On the other hand, for criticals of large magnitude, the shear rate is relatively high, and with the poor lubricating qualities of silicone oil as mentioned previously, the outer rubbing surfaces of the damper may gall, and eventually the damper mass will weld itself to the housing. Reference (4) lists only nickel and cadmium as satisfactory materials to act with steel when silicone is used as the lubricant. Accordingly, it has been necessary to cadmium-plate the damper flywheel in order to obtain a damper which must operate for long periods of time at a severe critical.

To assemble the damper, a coating of silicone oil is painted on

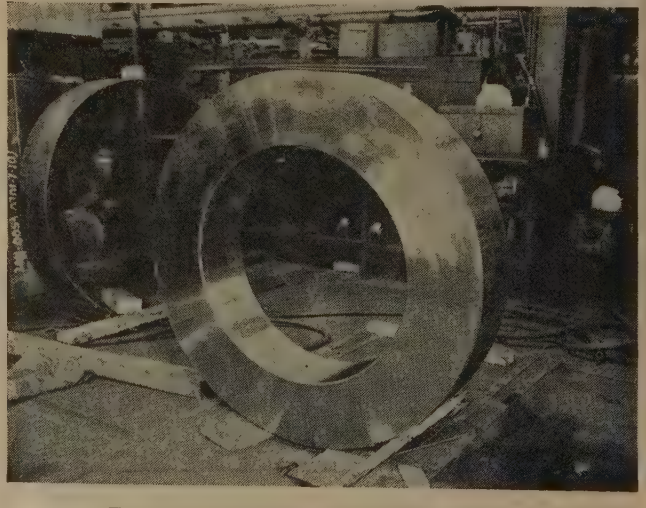


Fig. 5 Flywheel for 56-In-Diam Damper

the inside surfaces of the housing, and the inertia mass is carefully lowered into place. The cover plate is fitted to the housing and hermetically sealed by welding, except for two filling holes.

After welding, it becomes necessary to get the silicone fluid inside the housing. This is accomplished by first forcing the silicone oil by air pressure into the housing through one of the filling holes, shown as the two holes at the inner radius in Fig. 4. The silicone oil is placed in the filling bowl, and shop air is applied as shown in Fig. 6. The oil is forced in until it starts oozing out of the opposite filling hole. The filling equipment is then removed, the filling hole is plugged, and the damper assembly is spun at high speed, Fig. 7, forcing the silicone oil to the outer periphery of the damper. The filling and spinning procedure described is repeated until the damper is completely filled with the oil. As the silicone oil has low surface tension, it tends to leak through very small openings; hence it is necessary to seal the plugged filling holes with solder, otherwise the silicone oil will ooze out around the edges of the pipe plugs.

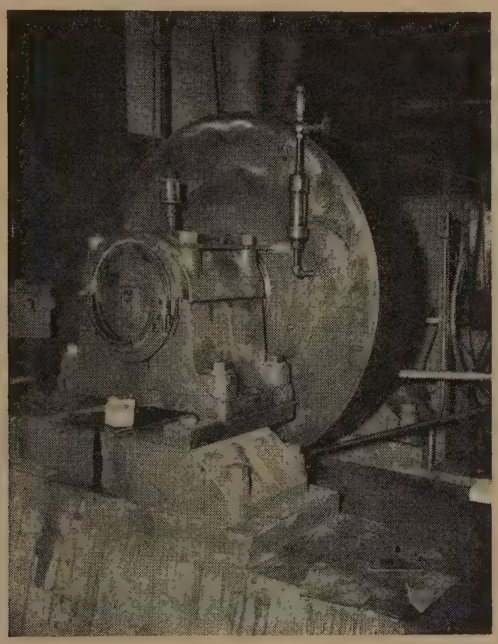


FIG. 6 FILLING 56-IN-DIAM DAMPER WITH SILICONE OIL

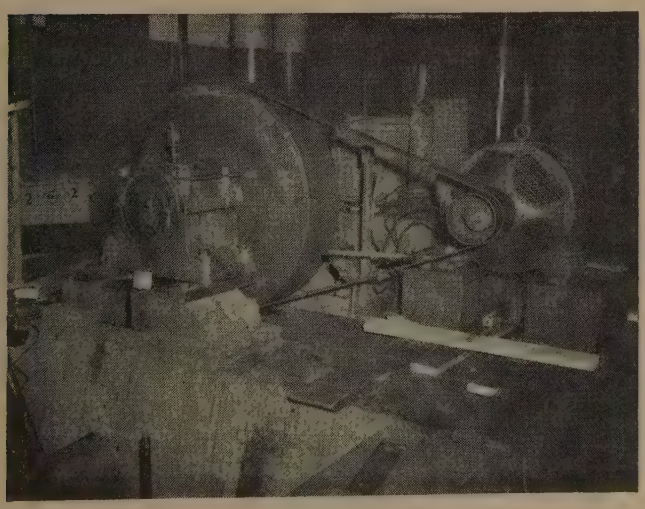


FIG. 7 SPINNING 56-IN-DIAM DAMPER

Fig. 8 shows a 32-in-diam damper mounted on the crankshaft of a 6-cylinder 16-in. × 22-in. supercharged four-cycle engine. This damper was manufactured by the Houde Engineering Division of Houdaille-Hershey Corporation.

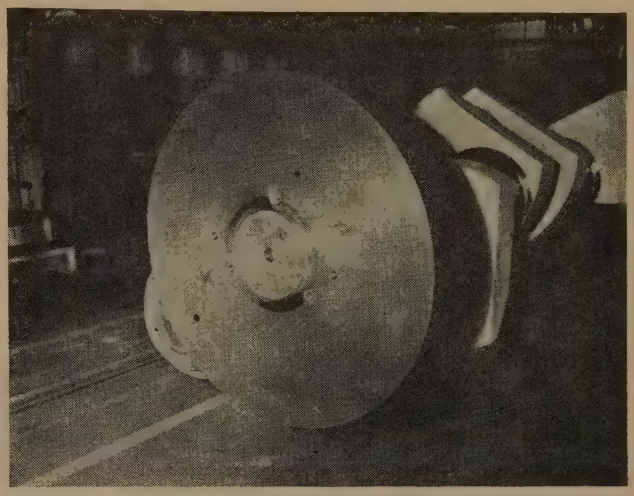


FIG. 8 32-IN-DIAM DAMPER MOUNTED ON CRANKSHAFT

TEST RESULTS OF DAMPER

In Table 2 are listed the sizes of engines with the calculated and test amplitude in degrees at the free end of the crankshaft for engines which have had dampers attached to them. The close agreement between the test and calculated results will be noted. Fig. 12 shows stress curves of a 29-in. × 40-in. marine engine with and without viscous damper.

TABLE 2 CALCULATED AND TEST AMPLITUDES FOR ENGINE WITH DAMPERS

Engine type	Critical order no.	Calculated amplitude, deg	Test amplitude, deg
6-Cylinder, 16 × 22, 4-cycle, supercharged, 1200 bhp	9	±0.18	±0.17
7-Cylinder, 16 × 22, 4-cycle, supercharged, 1400 bhp	7	±0.58	±0.46
8-Cylinder, 16 × 22, 4-cycle, supercharged, 1600 bhp	8	±0.57	±0.44
7-Cylinder, 16½ × 24½, 4-cycle, supercharged, 1575 bhp	7	±0.57	±0.43
5-Cylinder, 21½ × 29, 2-cycle, 1875 bhp	10	±0.15	±
5-Cylinder, 29 × 40, 2-cycle, 3100 bhp	10	±0.11	±0.07
10-Cylinder, 29 × 40, 2-cycle, 7100 bhp	7	±0.91	±1.14
	10	±0.94	±0.98

CONCLUSIONS

We have shown in this paper that the torsional viscous-friction damper may be readily designed by use of simple formulas for a given engine to obtain reduced stresses in the crankshaft. These formulas for a given engine will furnish complete information for the physical dimensions of the damper mass by a simple cut-and-try method, and give the thickness and viscosity of the oil film. It is only necessary to exercise care in the manufacture of the damper, in order to obtain a damper which will last for the life of an engine with little or no attention or maintenance.

ACKNOWLEDGMENT

The author wishes to acknowledge the helpful co-operation of the General Electric Company and Dow Corning Corporation in supplying him with the physical properties of the silicone oils manufactured by them.

BIBLIOGRAPHY

1 British Patent No. 21139, Lanchester patentee.

2 "The Elimination of Torsional Vibration," by G. J. Dashefsky, The Pennsylvania State College Bulletin—The School of Engineering, Technical Bulletin No. 12, Nov. 1, 1930, pp. 193–267.

3 "Dimethyl-Silicone-Polymer Fluids and Their Performance Characteristics in Unilaterally Loaded Journal Bearings," by J. E Brophy, R. O. Militz, and W. A. Zisman, Trans. ASME, vol. 68, 1946, pp. 355–360.

4 "Dimethyl-Silicone-Polymer Fluids and Their Performance Characteristics in Hydraulic Systems," by V. G. Fitzsimmons, D. L. Pickett, R. O. Militz, and W. A. Zisman, Trans. ASME, vol. 68, 1946, pp. 361–369.

5 "Silicone Oils," by D. F. Wilcock, *General Electric Review*, vol. 49, 1946, pp. 14–18 and 28–33.

6 "Practical Solution of Torsional Vibration Problems," by W. Ker Wilson, John Wiley & Sons, Inc., New York, N. Y., second edition, 1941.

7 "Mechanical Vibrations," by J. P. Den Hartog, McGraw-Hill Book Company, Inc., New York, N. Y., 1940.

8 "Evaluation of Effects of Torsional Vibration," SAE War Engineering Board, Society of Automotive Engineers, Inc., New York, N. Y., 1945, pp. 232–267.

9 "Lubrication," by A. E. Norton, McGraw-Hill Book Company, Inc., New York, N. Y., 1942.

10 "The Viscous Torsional Vibration Damper," by B. E. O'Connor, SAE Quarterly Transactions, vol. 1, January, 1947, pp. 87–97.

Appendix

Theory of Torsional Viscous-Friction Damper as Applied to Multimass System

Nomenclature

The following nomenclature is used in the Appendix:

I_d = moment of inertia of flywheel, lb-in-sec²

I_h = moment of inertia of housing, lb-in-sec²

$\Sigma I\theta^2$ = Effective moment of inertia of system referred to free end of crankshaft. This is the arithmetic sum of the the product of the individual masses and the square of its amplitude from the normal elastic curve or Holzer tabulation. (This includes the damper housing plus one half of the damper mass.), lb-in-sec²

c = damping coefficient, lb-in. per radian per sec

θ_d = absolute amplitude of flywheel, radian

θ_1 = absolute amplitude of housing, radian

θ_r = relative amplitude of flywheel, radian

ω = circular frequency, radian per sec

W_d = damping work of damper, in-lb per sec

W_e = input work of engine, in-lb per sec

W_{de} = damping work in engine, in-lb per sec

T_n = nth order harmonic coefficient, psi

A = area of piston, sq in.

R = crank radius, in.

R_o = outer radius of inertia flywheel, in.

R_i = inner radius of inertia flywheel, in.

μ = absolute viscosity, lb-sec per sq in. or reyn

h = oil-film thickness, in.

F = force acting on oil film, lb

a = area of oil film, sq in.

S = shearing force in oil film, psi

V = velocity, ips

Ω = angular velocity of flywheel, radian per sec

b = width of inertia flywheel, in.

γ = density of flywheel material, lb per cu in.

τ_{damp} = damped vibratory stress in crankshaft, psi

τ_{max} = vibratory stress in crankshaft with engine damping only, psi

Referring to Fig. 9, the differential equation of motion of the inertia flywheel is

$$I_d \frac{d^2\theta_d}{dt^2} + c\left(\frac{d\theta_d}{dt} - \frac{d\theta_1}{dt}\right) = 0 \quad \ldots\ldots [1]$$

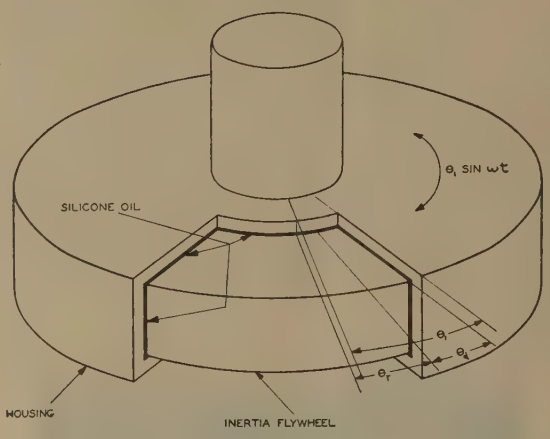


Fig. 9 Section of Inertia Flywheel and Housing

If the relative motion

$$\theta_r = \theta_d - \theta_1 \sin \omega t \quad \ldots\ldots [2]$$

is substituted in Equation [1], we obtain the differential equation of motion in terms of the relative displacement of the inertia flywheel or damper mass

$$I_d \frac{d^2\theta_r}{dt^2} + c\frac{d\theta_r}{dt} = I_d \theta_1 \omega^2 \sin \omega t \quad \ldots\ldots [3]$$

The particular solution of this differential equation may be shown to be

$$\theta_r = \frac{\theta_1 \sin(\omega t - \phi)}{\sqrt{1 + \left(\frac{c}{I_d \omega}\right)^2}} \quad \ldots\ldots [4]$$

where

$$\tan \phi = -\frac{c}{I_d \omega}$$

This equation gives us the relative motion of the inertia flywheel in terms of the absolute motion of the housing, moment of inertia of the inertia flywheel, circular frequency, and the damping coefficient. The only quantity that can be varied in Equation [4] is the damping coefficient for a given size of inertia flywheel. The optimum value of the damping coefficient must be such as to give maximum damping work from the damper. To find this optimum value the damping work is formulated from the fact that the damping force is 90 deg out of phase with the displacement; therefore the work dissipated per damping cycle (6, 7) is

$$W_d = \pi c \omega \theta_r^2 \quad \ldots\ldots [5]$$

Substituting Equation [4] in [5], the damping work in terms of the absolute motion of the housing is

$$W_d = \frac{\pi c \omega \theta_1^2}{1 + \left(\frac{c}{I_d \omega}\right)^2} \quad \ldots\ldots [6]$$

Differentiating the damping work with respect to the damping coefficient c, and equating the differential to 0, the optimum value of c is found to be

$$c = I_d \omega \quad \ldots\ldots [7]$$

Substituting Equation [7] in Equation [6], the maximum damping work is

$$W_d = \frac{\pi I_d \omega^2 \theta_1^2}{2} \quad \ldots\ldots [8]$$

All of the foregoing equations presuppose that the value of the natural circular frequency is known, though actually the value of ω depends upon the value of the damping coefficient. However, it will now be shown that the natural frequency of the multimass system may be calculated by using an equivalent inertia of the flywheel at the housing.

The differential equation of the motion of the housing is

$$I_h \frac{d^2\theta_1}{dt^2} + k(\theta_2 - \theta_1) + c\left(\frac{d\theta_1}{dt} - \frac{d\theta_d}{dt}\right) = 0 \quad \ldots\ldots [9]$$

where

k = stiffness of shaft between housing and next mass, usually cylinder No. 1, in-lb per radian

θ_2 = amplitude of next mass, usually cylinder No. 1, radians

Substituting Equation [1] in Equation [9], the differential equation for the housing becomes

$$I_h \frac{d^2\theta_1}{dt^2} + I_d \frac{d^2\theta_d}{dt^2} + k(\theta_2 - \theta_1) = 0 \quad \ldots\ldots [10]$$

The absolute acceleration of the inertia flywheel is equal to the sum of the absolute acceleration of the housing plus the relative acceleration of the inertia flywheel to the housing or

$$\frac{d^2\theta_d}{dt^2} = \frac{d^2\theta_1}{dt^2} + \frac{d^2\theta_r}{dt^2} \quad \ldots\ldots [11]$$

The relative acceleration is

$$\frac{d^2\theta_r}{dt^2} = \frac{\sin(\omega t - \phi)}{\sin \omega t \sqrt{1 + \left(\frac{c}{I_d \omega}\right)^2}} \frac{d^2\theta_1}{dt^2} \quad \ldots\ldots [12]$$

The average value of the relative acceleration for one cycle may be shown to be

$$\frac{d^2\theta_r}{dt^2} = -\frac{1}{1 + \left(\frac{c}{I_d \omega}\right)^2} \frac{d^2\theta_1}{dt^2} \quad \ldots\ldots [13]$$

Substituting Equation [13] in Equation [11], and then Equation [11] in Equation [10], the differential equation of motion becomes

$$\left\{ I_h + I_d \left[\frac{\left(\frac{c}{I_d \omega}\right)^2}{1 + \left(\frac{c}{I_d \omega}\right)^2} \right] \right\} \frac{d^2\theta_1}{dt^2} + k(\theta_2 - \theta_1) = 0 \quad \ldots [14]$$

Using the equivalent inertia

$$I_e = I_d \left[\frac{\left(\frac{c}{I_d \omega}\right)^2}{1 + \left(\frac{c}{I_d \omega}\right)^2} \right] \quad \ldots\ldots [15]$$

in Equation [14], the differential equation of the motion of the housing finally reduces to

$$(I_h + I_e) \frac{d^2\theta_1}{dt^2} + k(\theta_2 - \theta_1) = 0 \quad \ldots\ldots [16]$$

Thus to calculate the natural frequency of a multimass system with a viscous damper at a free end, as a conventional multimass system, it is only necessary to reduce the damper to an equivalent inertia at the free end by Equation [15]. For optimum damping, the equivalent inertia of the damper mass is equal to one half the actual inertia of the inertia flywheel.

It now becomes possible to calculate the amplitude at the free end of the engine with the damper attached. This is done by equating the input work done by the exciting forces in the engine to the damping work dissipated in the engine due to engine damping, and the work dissipated in the damper. The work done (6, 7) by the exciting forces is

$$W_e = \pi |T_n| \Sigma\theta \cdot \theta_1 \quad \ldots\ldots [17]$$

where

$|T_n|$ = nth order harmonic-torque component per cylinder, i.e., $(T_n \cdot A \cdot R)$. Excellent charts of the harmonic coefficients T_n may be found in reference (8).

$\Sigma\theta$ = vector sum, from normal elastic curve and phase diagrams, for nth order critical speed

The damping work lost in the engine may be calculated using several different methods. We shall use a method described by W. K. Wilson,[3] which is an empirical formula based on the assumption that the over-all damping loss in an engine is proportional to the cube of the vibratory stress in the crankshaft. The damping work lost in the engine is

$$W_{de} = \pi c_e \omega \Sigma\theta^2 \cdot \theta_1^2 \quad \ldots\ldots [18]$$

Where c_e is an equivalent engine-damping torque in inch-pounds per radian per second, and is found from the dynamic magnifier M, as follows[4]

$$M = \frac{250{,}000}{\tau_{\max}} = \frac{\omega \Sigma I\theta^2}{c_e \Sigma\theta^2}$$

or

$$c_e = \frac{\omega \Sigma I\theta^2}{\Sigma\theta^2} \frac{\tau_{\max}}{250{,}000} \quad \ldots\ldots [19]$$

Substituting Equation [19] in [18], we obtain

$$W_{de} = \frac{\pi \omega^2 \Sigma I\theta^2 \tau_{\max} \cdot \theta_1^2}{250{,}000} \quad \ldots\ldots [20]$$

The damping work of the damper is given by Equation [8]. Equating the work done by the exciting forces to the damping work, i.e., Equation [17] = Equation [8] + Equation [20] or

$$\pi |T_n| \Sigma\theta \cdot \theta_1 = \pi I_d \omega^2 \frac{\theta_1^2}{2} + \frac{\pi \omega^2 \Sigma I\theta^2 \tau_{\max} \cdot \theta_1^2}{250{,}000} \quad \ldots\ldots [21]$$

Solving for θ_1 we have

$$\theta_1 = \frac{|T_n| \Sigma\theta \cdot 250{,}000}{\omega^2 \Sigma I\theta^2 \tau_{\max}} \Big/ \left(1 + \frac{125{,}000\, I_d}{\tau_{\max} \Sigma I\theta^2}\right) \quad \ldots\ldots [22]$$

[3] Reference (6), vol. II, p. 118.

[4] Ibid., pp. 44 and 118.

but

$$\frac{|T_n|\Sigma\theta \cdot 250{,}000}{\omega^2 \Sigma I\theta^2 \tau_{max}} = \theta_e \quad [23]$$

is the amplitude at the forward end of the engine with engine damping only. Substituting Equation [23] in [22] we obtain

$$\theta_1 = \frac{\theta_e}{1 + \dfrac{125{,}000\, I_d}{\tau_{max}\, \Sigma I\theta^2}} \quad [24]$$

but $\theta_1\tau = \tau_{damp}$ and $\theta_e\tau = \tau_{max}$, where τ is the crankshaft stress in pounds per square inch per degree amplitude at the forward end of the crankshaft. Substituting these in Equation [24] we get

$$\tau_{damp} = \frac{\tau_{max}}{1 + \dfrac{125{,}000\, I_d}{\tau_{max}\, \Sigma I\theta^2}} \quad [25]$$

This equation was developed by using a given type of engine damping; for other types of engine damping similar formulas may be developed. All of these formulas will have the same characteristic that the damped vibratory stress will depend on the ratio of the moment of inertia of the damper mass to the effective moment of inertia of the vibratory system referred to the free end of the crankshaft. To reduce a given critical vibratory stress to some allowable value, it is only necessary to select such a size of the damper inertia that it will reduce the undamped vibratory stress to this allowable value by use of Equation [25]. It is to be noted, however, that a damper inertia larger than 50 per cent of the equivalent inertia will be of little effect in reducing vibratory stress to still lower values, as any addition of the damper inertia increases the equivalent inertia of the system, so all that is being done is to increase the damper inertia to damp itself.

Determination of Damping Coefficient c for Viscous-Friction Damper

The damper works on the basis that the energy is destroyed by the viscous friction between the damper mass and its housing. Referring to Fig. 10, Newton's law (9) for viscous friction is

$$S = \frac{F}{a} = \frac{\mu V}{h} \quad [26]$$

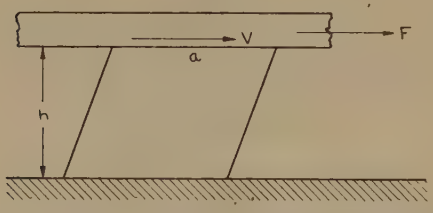


Fig. 10 Friction Block Illustrating Newton's Law

For rotation, Equation [26] becomes

$$SR = \frac{FR}{a} = \frac{\mu VR}{h} \quad [27]$$

The damping coefficient c is given by the expression

$$c = \frac{FR}{\Omega} = \frac{\mu R^2 a}{h} \quad [28]$$

For the viscous-friction damper, the over-all damping coefficient is the sum of the elements of the damping coefficient for each area of the damper, i.e.

$$c = c_i + c_o + c_s \quad [29]$$

For the inner surface of the damper we have

$$c_i = \frac{2\pi\mu b_i R_i^3}{h} \quad [30]$$

and for the outer surface

$$c_o = \frac{2\pi\mu b R_o^3}{h} \quad [31]$$

Referring to Fig. 11, the damping coefficient for the sides of the damper is

$$c_s = 2\int_{R_i}^{R_0} \frac{2\pi\mu r^3 dr}{h}$$

$$= \frac{\pi\mu}{h}(R_o^4 - R_i^4) \quad [32]$$

Substituting Equations [30], [31], and [32] in [29], the over-all damping coefficient is

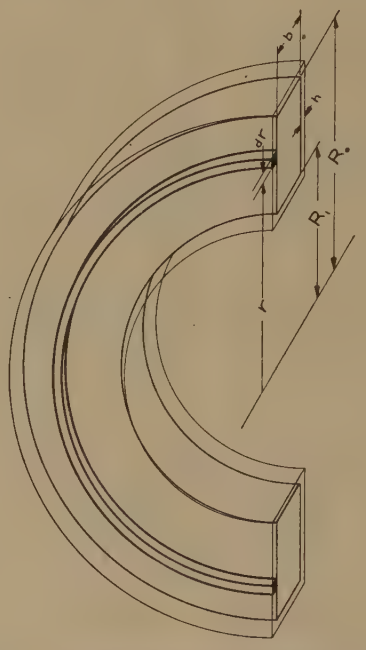


Fig. 11 Illustration of Sides of Damper

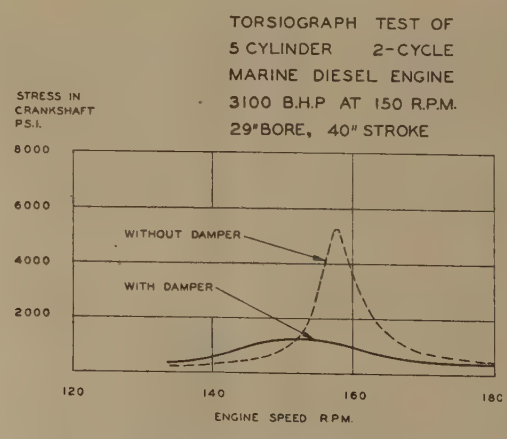


Fig. 12 Torsiograph Test of Marine Diesel Engine

$$c = \frac{\pi\mu}{h}\left[(R_o^4 - R_i^4) + 2b(R_o^3 + R_i^3)\right] \quad [33]$$

Substituting Equation [7], the optimum damping, in Equation [33] we have

$$I_d\omega = \frac{\pi\mu}{h}\left[(R_o^4 - R_i^4) + 2b(R_o^3 - R_i^3)\right] \quad [34]$$

but

$$I_d = \frac{\pi b\gamma}{2g}(R_o^4 - R_i^4) \quad [35]$$

and substituting Equation [35] in [34], and, after some algebraic manipulations we obtain

$$\mu = \frac{\frac{\gamma h\omega R_o}{4g}}{\frac{1+\left(\frac{R_i}{R_o}\right)^3}{1-\left(\frac{R_i}{R_o}\right)^4} + \frac{R_o}{2b}} \quad [36]$$

This equation gives the relation between the dimensions of the damper mass, the viscosity, and thickness of the oil film for the optimum damping coefficient c.

For further discussion of the theory of the viscous-friction damper see references in the Bibliography.[5]

As the silicone fluid used in the damper is actually a non-Newtonian fluid as the viscosity varies with shear rate as shown in Fig. 2, it is therefore necessary to know the value of the shear rate in the fluid for a given damper. The shear rate on the outer surface is

$$\text{Shear rate} = \frac{V}{h} \quad [37]$$

but $V = \theta_r\omega R_o$, and from Equation [4], for optimum damping, $\theta_r = \theta_1/\sqrt{2}$; substituting these in Equation [37] we have

$$\text{Shear rate} = \frac{\omega\theta_1 R_o}{\sqrt{2}\,h} \quad [38]$$

For a given design of damper all the factors in Equation [38] are known and therefore the shear rate can be determined, and the actual viscosity of the silicone oil also can be determined from Fig. 2. This value of viscosity must be at least larger than the value determined from Equation [36] for a properly designed damper.

Discussion

G. J. Dashefsky.[6] The author's contribution to the timely subject of viscous-friction dampers should prove to be of great practical value to designers who have to cope with the torsional-vibration problem. Particularly, his data regarding characteristics of silicone fluids used in these dampers and his detailed exposition of methods used for the filling of the dampers should prove to be of great interest. The information regarding lubricating qualities of the silicone oils and the necessity for coating the steel casing and flywheel properly are vital hints. The paper refers to work on viscous-friction dampers accomplished at the New York Naval Shipyard back in 1930. It is desired to mention here that much credit for the concept of the viscous-friction damper designed at this Yard should go to Capt. H. M. Jensen, U.S.N., who was then the Yard's Engineering Officer. At that time the materials available as a working fluid had poor temperature-viscosity characteristics, but, nevertheless, permitted operation of a tolerably useful device along the lines described in the author's reference (2).

In designing the early damper, consideration had to be given to the change of viscosity with rise in temperature resulting from the dissipation of vibrational energy by shearing of the viscous medium during the damping process, and conduction of heat to the damper casing through the crankshaft. In the viscous-friction damper, as in others of the Lanchester type, the optimum energy dissipation is related to a definite ratio between the exciting torque of the engine and the friction torque of the damper. Since the friction torque of the viscous-friction damper depends on the viscosity coefficient of the damper fluid, it is apparent that the optimum damping for a given damper will depend upon temperature. Hence there is a critical temperature for each damper application, at which the energy dissipation will be a maximum. For this reason the damper was divided into two compartments as shown in the author's reference (2), each having a distinct optimum temperature, both temperatures being so related by design so that the over-all damper characteristic was flattened out considerably with respect to temperature sensitivity. With the advent of the silicone oils which have a more suitable temperature-viscosity characteristic, as shown in the paper, the necessity for bicompartmental construction becomes unnecessary.

The mathematical solution for this type of damper has been presented previously in several places. Carter's exposition in reports of British Aeronautical Research Committee about 1926 or 1927, cover essentially the same ground as that in the present paper and in Mr. O'Connor's paper, reference (10) of the author. In the Appendix to the present paper the author presents an analysis which includes consideration of engine damping. In this discusser's opinion this unnecessarily complicates the calculations without sufficient compensation in accuracy of the results. Hysteresis damping falls off rapidly with decreasing stress. It is believed that a design which considers the damping contributed by the damper only will give sufficiently close results.

The viscous friction damper possesses several desirable characteristics. It is generally effective on several orders of vibration simultaneously, it is simple in mechanical construction, and quite satisfactory with regard to its weight when compared with other dampers.

The paper is a timely and important contribution which will be much appreciated by the engineer concerned with the torsional-vibration problem.

B. E. O'Connor.[7] When the viscous torsional-vibration damper, employing silicone in shear as a viscous damping medium, was introduced to the internal-combustion-engine industry by the writer in 1945, an extensive survey of the industry disclosed no use of viscous dampers on crankshafts. Since that time many manufacturers of gasoline and Diesel engines have adopted the Houdaille viscous damper for control of crankshaft torsional vibration. These developments are the subject of many patent applications by the writer.

The solution offered by the author will lead to reasonably accurate results, if the inherent engine damping for the engine under

[5] Reference (10); reference (6), vol. II, pp. 481–503; and reference (7), pp. 124, 129, and 253–258.

[6] Consultant Mechanical Engineer, Material Laboratory, New York Naval Shipyard, Brooklyn, N. Y. Mem. ASME.

Note: The opinions expressed in this discussion are those of the discusser and not necessarily those of the Navy Department nor the service at large.

[7] Houdaille Hershey Corporation, Houde Engineering Division, Buffalo, N. Y.

consideration follows the particular law which the author chooses to incorporate in his expression [25], and if the damper inertia is small compared with the inertia of the crankshaft. This method of solution was proposed by the writer in reference (10) of the author's Bibliography, and has been used satisfactorily with various inherent engine-damping expressions, in specifying dampers for many installations.

Since the solution as stated is at best an approximation, even if the nature of the inherent engine damping is known, it would appear prudent to disclose the nature of the approximation.

Equations [6], [7], and [8] as given by the author for the work per cycle, optimum damping, and maximum damping work, respectively, were derived by the writer in reference (10). It should be pointed out that the value of optimum damping derived in this manner does not take into account the effect of the damper on the frequency of the system and therefore does not represent optimum damping unless the damper mass is negligible in comparison with the other masses in the system. A somewhat lower value of damping will increase the frequency of maximum forced amplitude, thus increasing the work per cycle done by the damper.

The use of an equivalent inertia equal to the value of the damper-housing inertia plus one half the inertia of the damper flywheel was also proposed in reference (10). It is necessary, however, to question the validity of the author's derivation, since he uses $I_d\omega$ as the optimum damping whereas it can be shown, as follows, that the equivalent inertia of $I_h + I_d/2$ which leads to the proper value of the frequency of maximum forced amplitude, requires a value of optimum damping other than $I_d\omega$, except for the degenerate case of $I_d = 0$.

Equations [1] and [9] are identical to Equations [2] and [1] of reference (10), with the exception of the terms $K\theta_2$ and M_0. These terms also are seen to be identical since they both represent static deflection if divided by k.

The solution of this pair of equations, as given by Den Hartog in reference (7), for the frequency of maximum forced amplitude is, using the author's notation

$$\omega = \sqrt{\left(\frac{k}{I_h}\right)\left(\frac{2}{2 + \frac{I_d}{I_h}}\right)} \qquad [39]$$

for optimum damping, and the value of optimum damping is given as

$$c = \frac{2I_d\sqrt{k/I_h}}{\sqrt{2\left(1 + \frac{I_d}{I_h}\right)\left(2 + \frac{I_d}{I_h}\right)}} \qquad [40]$$

Equation [39] readily reduces to

$$\omega = \sqrt{\frac{k}{I_h + I_d/2}}$$

thus demonstrating rigorously that the frequency of maximum forced amplitude for the simple system under consideration is the same as though an inertia equal to one half the inertia of the damper flywheel had been added to the housing. It is to be recognized, however, that Equation [39] is based on the true value of optimum damping as expressed in Equation [40] and not on the value $I_d\omega$.

In terms of $I_d\omega$ Equation [40] reduces to

$$c = \frac{I_d\omega}{\sqrt{1 + \frac{I_d}{I_h}}} \qquad [41]$$

or expediently disregarding rigor

$$c = \frac{I_d\omega}{\sqrt{1 + \frac{I_d}{I_e'}}} \qquad [42]$$

for a multimass system where I_e' = the effective inertia of the entire system, less the inertia of the damper flywheel, referred to the location of the damper housing.

It is thus verified that the term $I_d\omega$ is a correct value of optimum damping only for the degenerate case of $I_d = 0$, and the accuracy of its use decreases as the size of the damper increases with respect to the inertia of the main system.

Equation [41] herewith, which is the optimum damping for the system which the author uses in his development, will give a value of c considerably different from $I_d\omega$ since I_d is large relative to I_h.

Equation [42] tends to justify the use of $I_d\omega$ as a damping constant for crankshaft installations since usually I_d is small compared to I_e'.

Fortunately, the value of damping is not critical. This property of the viscous damper as well as what the writer believes to be a simple justification of the use of $I_h + I_d/2$ as the equivalent inertia of the housing and flywheel in a Holzer table is given in reference (10).

The writer cannot agree with the author's statement, "a damper inertia larger than 50 per cent of the equivalent inertia will be of little effect in reducing vibratory stress to still lower values." Theory and practice both indicate otherwise.

Approximately true viscous damping presents many interesting possibilities other than the application to crankshafts. A self-damping acceleration-sensitive device, which responds to the same acceleration regardless of temperature, shimmy dampers for castoring wheels, and many other applications are at present the subjects of development and pending patent application.

Author's Closure

The author wishes to thank Mr. Dashefsky for his contribution of the history of his viscous-friction damper, and the reason for the bicompartmental construction. As he points out, this construction has now become unnecessary with the advent of the silicone oil.

The mathematical solution for this type of damper has been presented in several places; see references (11, 12, 13) in the bibliography at the end of this discussion. The author's contribution is in design calculation details.

Mr. Dashefsky objects to the use of engine damping in the author's analysis as he claims that this unnecessarily complicates calculations. While our design, Equation [25] is not a complicated form, it must be admitted that neglecting the engine damping increases the calculated damping stress about 20 per cent. It is believed, however, that the engine damping should be considered, as most of the test results in Table 2 of the paper show crankshaft stresses less than those determined from Equation [25].

The equations submitted by Mr. O'Connor, showing the nature of the author's approximate solution, are a welcome addition to the paper. It must be remembered that the exact solution for optimum damping for the usual size of dampers is of the order of 10 to 12 per cent of the approximate solution which is within the range of the manufacturing tolerances of the damper. There is not much point in calculating a damper more closely than it may be manufactured; also the approximate solution is on the safe side. As mentioned previously, the mathematics for this type of damper has been presented elsewhere, and, therefore, Mr. O'Connor's solution, as given in reference (10) of the paper, is not new.[8]

[8] Reference (6) of the paper, vol. 2, pp. 481–503.

Mr. O'Connor does not agree with the author's statement, "a damper inertia larger than 50 per cent of the equivalent inertia will by of little effect in reducing vibratory stress to still lower values." The author means by this statement that the addition of damper inertia beyond 50 per cent of the equivalent inertia will reduce stresses but not to an appreciable value,[9] e.g., doubling the damper inertia in the example in the paper reduces the calculated vibratory stresses 25 per cent.

[9] Reference (6) of the paper, vol. 2, p. 484.

BIBLIOGRAPHY

11 "Die Berechnung der Drehschwingungen," by H. Holzer, Julius Springer, Berlin, 1921.

12 "Verdrehungsschwingungen und ihre Dämpfung," by L. Gümbel, *Zietschrift des Vereines deutscher Ingenieure*, vol. 66, 1922, pp. 252–256 and 281–283.

13 "Torsional Vibration in Engines: Effects of Fitting a Damper, a Flywheel, or a Crankshaft Driven Supercharger," by B. C. Carter, Great Britain Aeronautical Research Committee, R & M 1053, Technical Reports 1926–1927, pp. 731–767.

Two-Species Laminated Beams

By A. G. H. DIETZ,[1] CAMBRIDGE, MASS.

This paper sets forth the simple theory of two-species wooden beams incorporating high-strength heavy outer laminations or "flanges," together with low-strength lighter cores into glued laminated members having high strength and high stiffness. Expressions for transformed section, rat'o of core depth to over-all depth, and shear at the neutral axis are derived. A few beams made of Douglas-fir outer laminations, and No. 1 or No. 2 common eastern-spruce cores were tested and their strength properties determined. Their performance coincided quite well with that predicted by the theory. A "form factor" similar to that for wooden I-beams may be involved.

General

GLUED laminated beams are customarily made of one species only. Furthermore, it is common practice to make such beams of only one grade of one species, although two grades are sometimes used, and rules respecting the use of two different strength grades in the same laminated beam have been worked out. Consequently, high-strength material sometimes is used on the outer edges of laminated beams and lower-strength material of the same species in the central portion or core. For example (1),[2] the material for the outer laminations of a beam may be selected so as to have knots not more than one half the size permitted in the grade to be used in the rest of the beam or core. When this selected material is placed on the outer edges of the beam it is up-graded to the next stress grade above that of the core, and the entire beam is considered to be designed on the basis of the up-graded outer laminations provided these are at least one tenth the thickness of the beam.

It is much less common practice to use two or more species of timber in the same glued laminated beam. It is the purpose of this paper to make an analysis of the possibilities in two-species beams in which a small amount of heavy strong material is used at the edges with larger quantities of lower-grade and different species in the core. It will be shown that relatively small amounts of the high-grade heavy species can be combined with relatively large amounts of the lower-grade lightweight species to yield beams having high strength. Designers of laminated beams might very well consider the advantages of using two different species as well as two different strength grades in the same laminated beam.

Relatively little has been done to analyze or test the possibilities and limitations of two-species versus one-species laminated beams. A. G. H. Dietz (2) published the results of a few tests of combined greenheart-spruce beams and made a brief analysis of the economics of such beams (3). H. McKean (4) made a more extensive study of laminated two-species beams. He used combinations of greenheart with redwood and white pine. In some instances his outer laminations or "flanges" were equal thickness, and in others the tension lamination was thinner. He found high-strength properties with relatively small amounts of greenheart. N. Poletika and F. Wangaard (5) have reported upon laminated beams of one species in which heavy strong material was placed at the edges and lighter weaker material in the core.

In this paper the simple theory of two-species beams is set forth and a few more test results are given.

Analysis of One-Species and Two-Species Beams

One-Species Beam. In Fig. 1 is shown a typical simply supported beam carrying several loads which induce a bending moment. Making the usual assumption in beam design that cross sections plane before bending remain plane after bending, the stress distribution across the depth of a beam such as this is as shown in Fig. 1(*b*). Since the material is assumed to be all of the same species and density even though it is glued laminated, the modulus of elasticity is the same throughout. Consequently the stress distribution across the depth of the beam is a straight line as shown in Fig. 1(*b*). If then the beam is composed of two

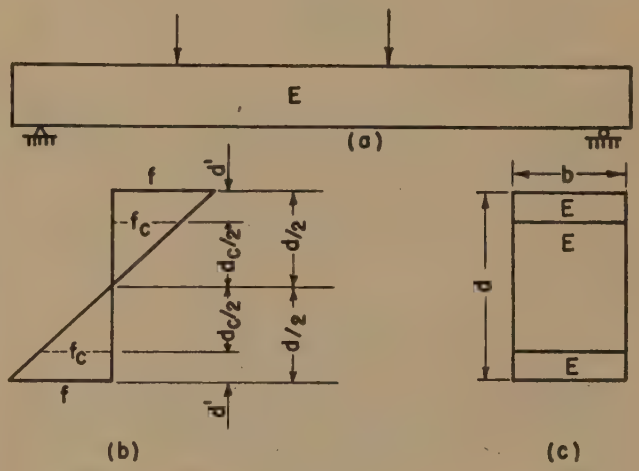


FIG. 1

different stress grades, a high-grade material of allowable stress f and a lower-grade material of allowable stress f_c, with the high-grade material used in the outer lamination and the lower-grade material throughout the depth of the core, the relationship of depth of core d_c to total depth d and stress f and f_c will be

$$\frac{f}{f_c} = \frac{d/2}{d_c/2} \quad \ldots\ldots [1]$$

or if

$$\frac{f}{f_c} = m \quad \ldots\ldots [2]$$

then

$$\frac{d_c}{d} = \frac{1}{m} = \frac{f_c}{f} \quad \ldots\ldots [3]$$

The moment of inertia I, is the same as for any plain rectangular beam

$$I = \frac{bd^3}{12} \quad \ldots\ldots [4]$$

[1] Associate Professor, Department of Building Engineering & Construction, Massachusetts Institute of Technology. Mem. ASME.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Wood Industries Division and presented at a meeting of the Wood Industries Division, High Point, N. C., October 14–15, 1948, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Paper No. 48—WD1-6.

Equation [3] shows the amount of lower-grade core material which should be used if both the core and the outer lamination are to be stressed to their maximum allowable fiber stresses. For example, if the value f of the higher-grade material is 1800 psi, and the value f_c of the lower-grade material in the core is 1200 psi, then two thirds of the entire beam will be the lower-grade material in the core, and one third will be the higher-grade in the edges, each of which is one sixth the total depth of the beam.

Two-Species Beam. In a two-species beam the heavier stronger species is placed in the outer lamination and the weaker lighter species in the core. The arrangement is as shown in Fig. 2(*a*). The outer laminations have a modulus of elasticity E_1 and the core has a modulus of elasticity E_2. If the usual assumption is made that cross sections plane before bending are plane after bending, the stress distribution across the depth of the beam is as shown in Fig. 2(*b*). The stress in the outermost fibers of the outer laminations is equal to f. At the interface between the outer lamination and the core, there is a break in the fiber stress distribution with a drop to f_c, as shown in Fig. 2(*b*). If, as before, the allowable stress in the outer laminations is f and the allowable stress in the core is f_c and if the ratio of moduli of elasticity E_1 and E_2 is

$$\frac{E_1}{E_2} = n \qquad [5]$$

then from Fig. 2(*b*)

$$\frac{f_c}{d_c/2} = \frac{f/n}{d/2} \qquad [6]$$

$$\frac{f}{f_c} = \frac{nd/2}{d_c/2} = m \qquad [7]$$

and

$$\frac{d_c}{d} = \frac{n}{m}, \quad \text{or} \quad d_c = \frac{n}{m}d = \frac{E_1}{E_2}\cdot\frac{f_c}{f}\cdot d \qquad [8]$$

The moment of inertia may be computed by employing a transformed section, that is, by multiplying the width b of the outer laminations by n, the ratio of the moduli of elasticity, in a manner analogous to that employed in reinforced-concrete design. The resulting transformed cross section is shown in Fig. 2(*c*). The transformed moment of inertia is

$$I_{tr} = \frac{nbd^3}{12} - \frac{(n-1)bd_c^3}{12} = \frac{b}{12}[nd^3 - (n-1)d_c^3] \qquad [9]$$

Substituting Equation [8] in Equation [9]

$$I_{tr} = \frac{bd^3}{12}\left[n - (n-1)\left(\frac{n}{m}\right)^3\right] = K\frac{bd^3}{12} \qquad [10]$$

For example, if as in the example of the one-species beam the allowable fiber stress f in the outer lamination is 1800 psi and the modulus of elasticity E_1 is 1,600,000 psi, and if the allowable fiber stress f_c of the core is 1200 psi and the modulus of elasticity E_2 is 1,400,000 psi, then d_c, the depth of the core, is 76 per cent of the total depth. This means that three quarters of the beam can be the lower-grade material in the core and only one quarter need be required for the edges, with each edge one eighth the total depth. This compares with the values of two thirds for the core and one third for the edges of the one-species beam. Therefore, in theory at least, a larger proportion of the two-species beam can be the lower-grade core material and still achieve a beam of the same bending strength. This is shown by comparing Equations [8] and [3].

The theoretical transformed moment of inertia given by Equation [10] is larger than the moment of inertia of the rectangular cross section by the amount shown in the bracket. The stiffness factor or EI value is based upon the modulus of elasticity of the core material. In the one-species beam the stiffness factor EI is of course based upon the same modulus of elasticity throughout the beam unless the outer laminations are stiffer than the core. In such an event, the two-species theory would apply to the one-species beam. From Equation [10] it is seen that if a beam is made of the two materials indicated, that is to say, materials having moduli of elasticity 1,600,000 psi in the outer laminations and 1,400,000 psi in the core with the respective fiber stresses of 1800 and 1200 psi, the transformed moment of inertia according to Equation [10] is 1.08 times as great as the moment of inertia of a plain rectangular beam of the same over-all dimensions. To obtain the stiffness factor, this transformed moment of inertia would be multiplied by the modulus of elasticity of the core, 1,400,000 psi. A solid beam all of modulus of elasticity 1,600,000

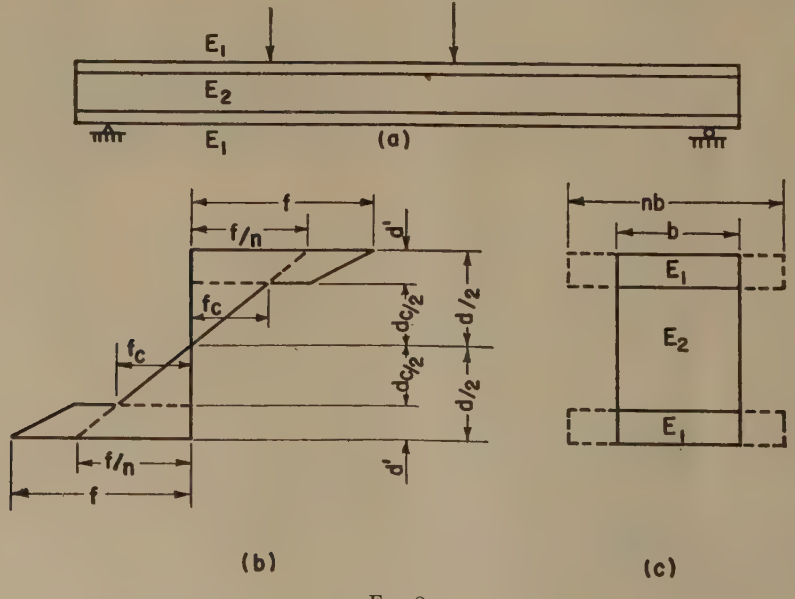


Fig. 2

psi on the other hand would have a higher EI value, even though the moment of inertia in that case would be merely that of the rectangular beam.

The horizontal shearing-stress distribution across the depth of of the transformed section or two-species beam is somewhat similar to that of an I-beam. In a plain rectangular beam, the shear distribution follows the usual parabolic distribution with a maximum at the neutral axis equal to 1.5 times the average. In a two-species beam the distribution is somewhat flatter across the core as would be expected from the analysis of the transformed section.

Horizontal shearing-stress intensity at any point is computed from the following expression

$$S = \frac{VQ}{bI} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [11]$$

For rectangular sections, at the neutral axis

$$S = \frac{3}{2} \cdot \frac{V}{bd} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [12]$$

From Equation [11] and the transformed section of Fig. 2(c), the horizontal shearing-stress intensity at the neutral axis of the two-species beam is found to be

$$S_{tr} = \frac{3}{2} \cdot \frac{V}{bd} \cdot \frac{m[m^2 - n(n-1)]}{m^3 - n^2(n-1)} \quad \ldots\ldots\ldots [13]$$

Using the same values of f, f_c, E_1, and E_2 as before, Equation [13] becomes

$$S_{tr} = 0.98 \times \frac{3}{2} \cdot \frac{V}{bd}$$

that is, the intensity of horizontal shearing stress at the neutral axis is slightly less than it would be in a plain beam of the same cross section. Depending upon the relative ratios of m and n, the reduction in shearing-stress intensity may range from a practical maximum of about 10 per cent down to zero. In relatively short heavily loaded beams, therefore, in which horizontal shear at the neutral axis may be the controlling factor, there may be a small advantage in the two-species beam.

The factor K in Equation [10] may be found from curves such as those in Fig. 4. Here, for example, for a value of m equal to 2, and n equal to 1.5, K is found to be 1.28.

Possible Limiting Factor in Two-Species Beams. The foregoing analysis assumes that the two-species beam will behave in bending in a similar manner to a single species. It has been pointed out, however, by J. A. Newlin and G. W. Trayer (6) and reported elsewhere as well (7, 8) that in a wooden I-beam or box beam, for example, this straightforward bending action does not occur. It is well known of course that the direct compressive strength of wood parallel to the grain is less than the computed modulus of rupture of the same material when placed in bending (6, 8, 9, 10), and that consequently form factors must be introduced in the design of I-beams which reduce the permissible bending stress. It may be that in a two-species beam there is enough action similar to that of an I-beam to require a reduction in the alllowable bending stresses in the outer lamination. Whether this is true can be determined only by test.

Experimental Work

To check the validity of the theoretical approach outlined, several sets of test beams were made incorporating eastern-spruce cores and Demerara-greenheart or Douglas-fir outer laminations. The spruce-greenheart combinations have already been reported (2) and the results will not be repeated here except to indicate that this series of tests showed that with materials so widely different as eastern spruce (No. 1 common) and greenheart, the laminated beams were substantially as strong as the solid greenheart, even though the greenheart represented less than one third the total. In those beams some lower (tension) laminations were the same thickness as the upper (compression) laminations, and in others the tension laminations were of the order of one half to two thirds the thickness of the compression laminations. In all instances the strength properties were substantially the same. In the much more extensive series of tests made later by McKean (4) with greenheart combined with either redwood or white pine, the same general results were obtained.

In a second series of tests Douglas fir carefully selected for straight grain, medium growth, and fairly high percentage of summer wood (estimated at 25–30 per cent) was combined with No. 1 and No. 2 common eastern spruce.

Preliminary tests were made on small clear samples $^3/_4$ in. deep $\times$ $1^1/_2$ in. wide $\times$ 12 in. long of both Douglas fir and spruce, cut from the same pieces subsequently used in the laminated beams, to determine the moduli of elasticity and to determine the bending strength of the Douglas fir. Results of these tests were as follows:

E_f (Douglas fir)	=	1,920,000 psi
E_s (spruce)	=	1,440,000 psi
f_f (Douglas fir)	=	11,700 psi (modulus of rupture)
	=	8100 psi (proportional limit)

Because the spruce was No. 1 and No. 2 common—not a structural grade—and therefore was expected to be quite variable in strength, an arbitrarily chosen maximum bending strength of f_s = 5800 psi was assigned. The tests later showed that this estimate was probably quite close to the correct average value for the spruce.

From the foregoing data, the ratio n of the moduli of elasticity was found to be

$$n = \frac{E_f}{E_s} = \frac{1{,}920{,}000}{1{,}440{,}000} = 1.33$$

and the ratio of bending stresses

$$m = \frac{f_f}{f_s} = \frac{11{,}700}{5800} = 2$$

From these values, the proportions of the beams were chosen as

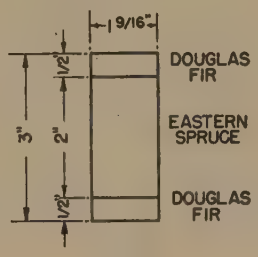


Fig. 3

shown in Fig. 3. With an over-all depth of 3 in., the depth of core was calculated

$$\frac{n}{m} = \frac{d_c}{d} = \frac{1.33}{2.00} = {}^2/_3$$

The spruce core, therefore, was 2 in. deep and the outer Douglas-fir laminations each $^1/_2$ in. thick. Glue was casein. Moisture content of wood measured with an electrical meter was 12 per cent.

All beams were simply supported at the ends and loaded at the

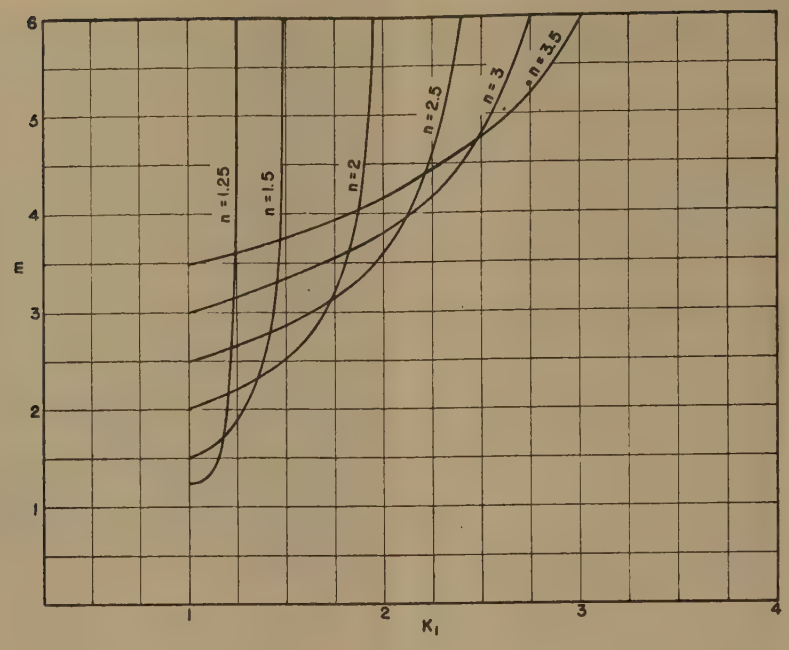


FIG. 4

center. Three beams were 34 in. span and four were 42 in. All beams were loaded to failure, and deflection readings were taken at the centers of the 42-in. beams. Maximum outer fiber stresses were computed for each beam in accordance with the computed transformed moment of inertia

$$f_u = n \frac{Mc}{I_{tr}} \quad [14]$$

EI values were first computed from the transformed sections and checked from the deflections of the 42-in. beams

$$EI = \frac{WL^3}{48d} \quad [15]$$

where d = deflection at center of beam. The results are summarized in Table 1.

TABLE 1 TEST RESULTS

E_f = 1,920,000 psi f_f = 11,700 psi (minor test specimen)
E_s = 1,440,000 psi f_s = 5800 psi
n = 1.33 m = 2

$$I_{tr} = \frac{bd^3}{12}\left[n - (n-1)\left(\frac{n}{m}\right)^3\right] = 4.32 \text{ in.}^4$$

Beam	Span, in.	Deflection at 1000-lb load, in.	Maximum load, lb.	*EI* Computed lb-in.^{2}	*EI* Measured, lb-in.^{2}	Maximum bending stress, psi
1	34	...	2840	...	...	11200
2	34	...	3000	...	...	11800
3	34	...	3090	...	...	12200
Average						11700
4	42	0.238	2250	6.22×10^6	6.50×10^6	11000
5	42	0.249	1800	6.22×10^6	6.20×10^6	8800
6	42	0.251	2200	6.22×10^6	6.15×10^6	10600
7	42	0.249	2000	6.22×10^6	6.20×10^6	9700
Averages				6.22×10^6	6.26×10^6	10000

It will be seen that the average maximum bending stress in the first three beams is the same as the average modulus of rupture of Douglas fir determined on the minor test specimens. In the second four beams, the average maximum bending stress is 10,000 psi or 86 per cent of that determined on the minor test specimens. In these four beams the measured average stiffness factor *EI* is almost exactly the same as the computed value.

In beam 1 the first signs of failure were compression wrinkles in the spruce, and in beam 2 the compression wrinkles began simultaneously in the spruce and the Douglas fir. In the balance of the beams the compression wrinkles all began in the fir but quickly spread into the spruce, or began at different points in the spruce soon after wrinkling began in the Douglas fir. These observed phenomena seemed to indicate that a close balance had been achieved between spruce and Douglas fir.

Conclusions and Recommendations

These and the previously reported spruce-greenheart tests are too few in number to permit any hard and fast conclusions to be drawn. They seem to indicate that two-species beams can be made with relatively large amounts of low-grade material of one species, combined with relatively small amounts of another higher-grade species to achieve beams substantially as strong and stiff as solid beams of the higher-strength higher-grade species. The quantity of high-grade material required appears, on the basis of the simple theory, to be less when two species—or to be more precise, two materials of different modulus of elasticity—are employed than if the high and low grades are all of the same species, or, more precisely, the same modulus of elasticity.

The test results bear out this theory in part. Stiffness seems to be as predicted by the theory, but the maximum bending stress in one set of beams is less than the minors, although in the other set the beams and the minors are the same. It may be that some sort of form factor is required here.

Much more testing will be required to establish the validity of the simple theory, but it is believed certain that high efficiency can be obtained with small quantities of high-grade species combined with large quantities of low-grade species.

In view of the fact that every possible economy must be achieved (3) if laminated beams, in spite of their many advantages, are to compete successfully with solid beams or with other structural materials, the possibilities in two-species beams should be carefully investigated by engineers. Furthermore, such beams open up possibilities in wood utilization with mixed species and mixed grades which may be attractive.

BIBLIOGRAPHY

1 "National Design Specifications for Stress-Grade Lumber and Its Fastenings," National Lumber Manufacturers Association, Washington, D. C., 1944 (with 1947 amendments), p. 52.

2 "Laminated Spruce and Greenheart Beams," by A. G. H. Dietz, *Tropical Woods*, Yale University School of Forestry, no. 56, 1938, pp. 1–5.

3 "High Grade and Low Grade Lumber in Glued Laminated Beams," by A. G. H. Dietz, Northeastern Wood Utilization Council, Bulletin No. 22, April, 1948.

4 "Mechanical Properties and Design Procedure for Glued Laminated Beams Composed of Two Wood Species," by H. B. McKean, Bulletin No. 11, School of Forestry & Conservation, University of Michigan, Ann Arbor, Mich., 1944.

5 "Laminated Connecticut Red Oak," by N. Poletika and F. Wangaard, Northeastern Wood Utilization Council Bulletin No. 22, April, 1948.

6 "Form Factors of Beams Subjected to Transverse Loading Only," by J. A. Newlin and G. W. Trayer, Report No. 181, National Advisory Committee for Aeronautics, Washington, D. C., 1924.

7 "Technical Data on Plywood," Section 7, Douglas Fir Plywood Association, Tacoma, Wash.

8 "Wood Handbook," Forest Products Laboratory, United States Department of Agriculture, Superintendent of Documents, Washington, D. C., 1940, pp. 154–155.

9 "Stress-Strain Relations in Timber Beams (Douglas Fir)," by A. G. H. Dietz, ASTM Bulletin, October, 1942.

10 "Strength and Related Properties of Woods Grown in the United States," by L. J. Markwardt and T. R. C. Wilson, Technical Bulletin No. 479, U. S. Department of Agriculture, Washington, D. C., 1935.

Appendix

In the final analysis, economic factors will usually dictate the use of one structural material in preference to another. It has seemed desirable, therefore, to attempt a cost analysis showing how costs of glued laminated beams might vary, depending upon whether they are composed of the following:

Case 1 All high-grade stock of fiber stress f and modulus of elasticity E.

Case 2 All one species (and modulus of elasticity E, but core of value $f_c = {}^2/_3 f$. Then $d_c = {}^2/_3 d$.

Case 3 Two species, modulus of elasticity of outer laminations E_1 is ${}^4/_3$ as great as modulus of elasticity E_2 of the core. Allowable fiber stress in outer laminations is f; allowable fiber stress in core $f_c = {}^2/_3 f$. Then $d_c = {}^8/_9 d$.

In each case it is assumed that the planing losses are such that 1.2 fbm of raw lumber will be needed to produce 1 fbm of laminated timber. Scarfing losses are assumed at 5 per cent in the outer laminations but only 3 per cent in the core laminations because butt joints can be used near the neutral axis.

In calculating the cost of the finished laminated beam, account must be taken not only of the cost of raw lumber, planing losses, and scarfing losses but, in addition, cost of glue per board foot (variable, depending upon glue used and thickness of laminations) labor, overhead, and profit. These can be put in the form of a cost equation as follows

$$x = b(1 + m)(y + d + e + f)(1 + p) + c(1 + n)(z + d + e + f)(1 + p) \quad [16]$$

$$= 1.155by + 0.127b + 1.133cz + 0.125c$$

If x = cost per board foot of laminated beam
y = raw material cost per board foot of the edge material, assumed \$60 per M
z = raw material cost per board foot of the core material, assumed \$30 per M
m = scarfing loss in edge material, assumed 5 per cent
n = scarfing loss in core material, assumed 3 per cent
b = quantity of edge material per board foot = $1.2\,(d - d_c)$
c = quantity of core material per board foot = $1.2\,d_c$
d = cost per board foot of glue, assumed \$10 per M
e = cost per board foot of labor, assumed \$40 per M
f = cost per board foot of overhead, assumed \$60 per M
p = profit per board foot, assumed 10 per cent

For the three cases listed, costs work out as follows:

Case 1 Suppose all of beam to be made of high-grade stock. Planing losses such that 1.2 fbm/bd ft needed, $b = c = 0.6$

$$x = 1.155 \times 0.6 \times \$0.06 + \$0.127 \times 0.6 + 1.133 \times 0.6 \times \$0.06 + \$0.125 \times 0.6$$
$$= 0.042 + 0.076 + 0.041 + 0.075$$
$$= \$234 \text{ per M}$$

Case 2 All same modulus of elasticity, but $f_c = {}^2/_3 f$, then $d_c = {}^2/_3 d$, and $b = 0.4$, $c = 0.8$

$$x = 1.155 \times 0.4 \times \$0.06 + \$0.127 \times 0.4 + 1.133 \times 0.8 \times \$0.03 + \$0.125 \times 0.8$$
$$= 0.028 + 0.051 + 0.027 + 0.100$$
$$= \$206 \text{ per M}$$

Case 3 $E_1 = {}^4/_3 E_2$, $f_c = {}^2/_3 f$, then $d_c = {}^8/_9 d$, and $b = {}^1/_9 \times 1.2 = 0.133$, $c = 1.067$

$$x = 1.155 \times 0.133 \times \$0.06 + \$0.127 \times 0.133 + 1.133 \times 1.067 \times \$0.03 + \$0.125 \times 1.067$$
$$= \$195 \text{ per M}$$

These costs appear to be in line with present prices for laminated timber such as are envisaged in this example. It is seen that by mixing high grades and low grades of the same species, the cost is reduced from \$234 to \$206 or about 12 per cent. By using mixed species, the cost is reduced still further to \$195 or a total of about 17 per cent. Obviously, all of these figures depend upon the relative raw-material costs of edge and core stock and all the other factors listed. Laminating as practiced today, however, is expensive in any event and advantage must be taken of every opportunity to reduce costs. Combining different grades, and combining different grades and species would seem to be one way to reduce those costs.

A Classification of Linear Transfer Members

By H. L. MASON,[1] AMES, IOWA

Linear dynamic behaviors are here grouped in a classification consisting of four basic curves with three methods of modification and two methods of combination. It depends on responses to unit step disturbance, and relates to Laplace transforms. Its utility is shown for members sequenced, manifolded, or looped, also for arbitrary disturbances and for reflex action.

Introduction

SCIENCE advances . . . by the continuous development of new and fruitful concepts," as J. B. Conant said recently in his presidential address to the AAAS. Those of us who are trying to crystallize into a science the art of designing and operating automatic controls are faced with the need of a set of basic concepts which shall describe the components which make up such dynamical systems. These concepts must be exact, so that there can be no ambiguity in identifying dynamic behavior; must be universal, so that identical properties can be recognized whether the member is found in a distillation process, a pneumatic valve, or an electrical thermometer; must be compact, so that they will be convenient in everyday use; must be familiar, so that they can be grasped readily by the mechanic, the college student, the practicing engineer, and the theoretician.

It is the purpose of this paper to meet these requirements with a systematic and comprehensive description of the behaviors of linear transfer members. The term "dynamical system" is here applied to any interconnected grouping of physical apparatus whose behavior is describable by time-dependent variables, whether the nature of the energy transfer be translational, rotational, hydraulic, pneumatic, thermal, electromagnetic, electrical, electronic, or chemical. We confine ourselves to cause and effect relations along one-dimensional network "paths," rather than in a two-dimensional field or in multidimensional space. A "disturbance" imposed at some point of the path, and proceeding from cause to effect, is imagined as exciting a consistent "response" in an adjoining variable. The action between them is said to be "reflexed" or "bilateral" if the responding variable can exert a reflexive effect on the disturbing variable; "cascading" or "unilateral" if it cannot. A path segment completely bounded by cascading actions is called a "transfer member." The member is "linear" if several forms of disturbance acting at once produce a response equivalent to the sum of their separate responses.

We choose as the basis for classifying linear transfer members the Cartesian curve pattern, plotted on a time base, of its response to a standard disturbance. This disturbance is shown in Fig. 1, and is called the "unit offset" or "unit step" at $t = 0$, in obvious description of its relation to the time axis. It will be designated $x_1(t)$, read "x-one of t," and defined by the equations $x_1 = 0$ for $t < 0$, $x_1 = 1$ for $t > 0$. It can be shown by Fourier analysis to comprise sinusoidal oscillations of all possible frequencies. Cascading action is assumed above it and below the response. The intervening member is assumed to be in equilibrium at $t = 0$; this implies that during some interval all varying quantities are temporarily constant, providing datum levels from which to measure the changes x, y, etc. It calls for zero values at $t = 0$ of all variables, of all derivatives except the highest, and of either the kinetic or the potential energies.

Fig. 1 Unit Offset

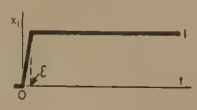


Fig. 2 Approximate Unit Offset

A response to $x_1(t)$ will be called an "indicial response," and designated $y_1(t)$. Some typical forms are shown in Figs. 3, 4, and 5. An indicial response provides an unambiguous description of the dynamic behavior of the associated transfer member, and from it can be predicted the behavior $y(t)$ of that member under any arbitrary disturbance pattern $x(t)$.

Both x_1 and y_1 patterns can be recognized and understood by anyone with a bit of mechanical sense. Although the disturbance is very abrupt, it can be practically realized in many physical systems. In others, it can be approximated, as by the steep ramp of Fig. 2, where $x_1 = 0$ for $t < 0$, t/ϵ for $0 < t < \epsilon$, and 1 for $t > \epsilon$, with $\epsilon << 1$. An indicial response can be found experimentally as the record on an instrument chart, graphically as a step-by-step construction, or analytically as the solution of an equation. It is thus equally useful to the mathematical and the nonmathematical thinker.

Four Basic Patterns

The indicial response, then, will be the dog-tag which identifies the transfer member from a dynamic point of view. Although we may encounter in automatic-control problems a well-nigh infinite variety of such responses, it appears that any of them which occurs in a linear segment or network with constant parameters can be classified as (i) one of four "basic" patterns, or (ii) one of three "modifications" of these, or (iii) two or more of the basic or modified patterns, joined by two "combining" methods. The four basic patterns for y_1 are shown in Fig. 3. Various names for these curves have been used in the literature; the adjectives chosen here are compromises between brevity, vividness, and exactness. "Offset" indicates a sudden and sustained shift relative to the time axis, "tardigrade" implies slow-moving in comparison with the originating disturbance, "sinusoid" describes a resemblance to the sine curve, "damping" infers a sinusoidal oscillation which is gradually damped out.

The mathematical expressions for these four indicial responses are listed in Table 1, together with a related symbol which we call the "transference," or transfer function. This will be defined later. For the moment it serves merely as a convenient code which defines the pattern concisely, and states the constants which fix the proportions of the graphical representation. Factor A is any real number, positive or negative, up to the value at which linearity no longer holds; a, δ, and ω are positive real numbers. For the offset, A is the sustained magnitude. For the tardigrade, A is the slope of y_1 at $t = 0$, and $1/a$ is the familiar time of 63 per cent response, which measures the slowness with which y_1 approaches the horizontal asymptote A/a. For the offset sinusoid, A/ω^2 is both the amplitude of the oscillation whose period

---

[1] Research Professor of Mechanical Engineering, Iowa State College. Mem. ASME.

Contributed by the Industrial Instruments and Regulators Division and presented at the Annual Meeting, New York, N. Y., November 28–December 3, 1948, of The American Society of Mechanical Engineers.

TABLE 1 BASIC PATTERNS

Name	Time Curve	Transference
(1) Offset.....	$Ax_1(t)$	A
(2) Tardigrade...	$\frac{A}{a}(1 - e^{-at})$	$\frac{A}{s+a}$
(3) Offset sinusoid	$\frac{A}{\omega^2}(1 - \cos \omega t)$	$\frac{A}{s^2+\omega^2}$
(4) Offset damping	$\frac{A}{\delta^2+\omega^2} - \frac{Ae^{-\delta t}\cos(\omega t - \tan^{-1}\delta/\omega)}{\omega(\delta^2+\omega^2)^{1/2}}$	$\frac{A}{(s+\delta)^2+\omega^2}$

is $2\pi/\omega$, and its mean value. For the offset damping, $A/(\delta^2 + \omega^2)$ is the steady value ultimately reached by the oscillation whose period is $2\pi/\omega$ and whose envelope curves decay exponentially. This last pattern evidently reduces to the offset sinusoid when $\delta = 0$; the unstable case of negative δ is not listed because it is an unwanted form.

We might choose other forms as basic, but these are the ones met most frequently and in relatively simple apparatus. On the other hand, much elaboration and complexity of apparatus may be necessary to produce these same patterns giving desired values of A, a, δ, or ω.

Three Modifications

The variety of indicial responses presented by useful groupings of apparatus is greater than the four basic types just listed. The

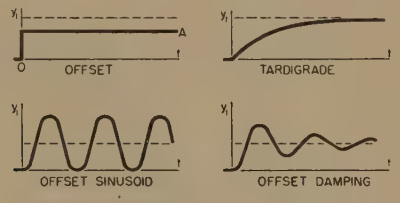


Fig. 3 Basic Indicial Responses

three following modifications of these patterns will extend their coverage:

1 Differentiation, giving the slope of the basic curve.
2 Integration, giving the area under the basic curve.
3 Delay, giving a shift of the basic curve to the right.

Fig. 4 shows some of the new patterns obtained by these modifications. Again, the names suggested are descriptive of the graphs; they will be obvious with the possible exception of "retraction," which expresses a sudden outthrust followed by a gradual withdrawal.

Let us now digress briefly to examine the meaning of the expression which we have called transference, and shall designate[2] by the Greek Λ. When the physical relationships are known, it is possible to write an equation relating disturbance and response through multiples, derivatives, integrals, differences, trigonometric forms, or hyperbolic forms. The individual terms of such an equation, which will be constants or functions of time, are transformed to functions of the complex variable s by the Laplacian transformation

$$\mathcal{L}[f(t)] = \int_0^\infty f e^{-st} dt = \tilde{f}(s)$$

where s is often written $\delta + j\omega$, and $j = \sqrt{-1}$. This gives an algebraic expression in terms of two Laplace transforms, $\tilde{y}(s)$ for a response and $\tilde{x}(s)$ for the disturbance which originates it. Transference Λ is defined as the ratio $\tilde{y}/\tilde{x}$, which for unit step disturbance x_1 becomes simply $s\tilde{y}_1$. The inverse transformation is given as

$$\mathcal{L}^{-1}[\tilde{f}(s)] = \frac{1}{2\pi j}\int_{-j\infty}^{+j\infty} \tilde{f} e^{ts}\, ds = f(t)$$

[2] Refer to note at end of author's closure, p. 412.

and both direct and inverse transforms exist for most of the functions met in automatic controls. Tables of function pairs $f(t)$ and $\tilde{f}(s)$ are available, and if necessary can be used much as is the phrase book of a foreign language, to translate expressions in the familiar "time" domain to those in the strange "complex" domain, and back again.

It may be noted that any specific transfer member has as many indicial responses, and transferences, as it has possible effect/cause relations. Consider, for example, a valve discharging a compressible fluid, without change of state, from a pressure $\bar{p}_a$ to a pressure $\bar{p}_b$ greater than critical. Small changes of pressure p_a or p_b or of lift x each causes a different flow change q, and there are three (approximately independent) transferences $\Lambda(\tilde{q}_a:\tilde{p}_a)$, $\Lambda(\tilde{q}_b:\tilde{p}_b)$, and $\Lambda(\tilde{q}_x:\tilde{x})$.

Under the rules of Laplace transformation, and with the initial equilibrium called for by our definition of indicial response, $\mathcal{L}(dy_1/dt) = s\tilde{y}_1$, so that modifying a basic response by differentiation becomes merely a multiplication of its transform by s. Similarly, integration corresponds to multiplication by $1/s$, and

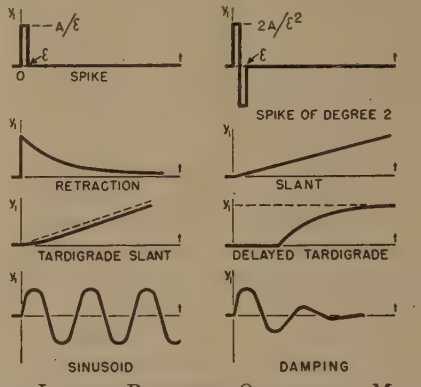


Fig. 4 Some Indicial Responses Obtained by Modification

delay of an interval b to multiplication by e^{-bs}. Basic indicial response $Ax_1(t)$, the "offset" with transference $\Lambda(\tilde{y}_1:\tilde{x}_1) = A$, may be approximated by a steep ramp with slope A/ϵ, resembling Fig. 2. Upon differentiating this piecewise, and letting $\epsilon \to 0$, a new indicial response is found as

$$\lim_{\epsilon \to 0} \frac{A[x_1(t) - x_1(t-\epsilon)]}{\epsilon} = Ax_1'(t)$$

which is the "spike" or "impulse" of Fig. 4, having $\Lambda = As$. Again, $Ax_1'(t)$ may be approximated by steep slopes of $2A/\epsilon^2$ and $-2A/\epsilon^2$. Upon differentiating this piecewise, and letting $\epsilon \to 0$, another indicial response is obtained as

$$\lim_{\epsilon \to 0} \frac{A[x_1(t) - 2x_1(t-\epsilon/2) + x_1(t-\epsilon)]}{\epsilon^2}$$

This is the "spike of degree 2," or "doublet impulse" $Ax_1''(t)$, having $\Lambda = As^2$. Upon one integration, the response $Ax_1(t)$ with $\Lambda = A$ becomes a new response with $y_1 = At$ and $\Lambda = A/s$; upon two integrations the "parabola of degree 2," $y_1 = At^2/2$ with $\Lambda = A/s^2$; upon a delay of interval b, the "delayed offset" $y_1 = Ax_1(t-b)$ with $\Lambda = Ae^{-bs}$.

Basic indicial response $y_1 = (A/a)(1 - e^{-at})$, the "tardigrade," with $\Lambda = A/(s+a)$, becomes upon one differentiation the "retraction" $y_1 = Ae^{-at}$ with $\Lambda = (As)/(s+a)$; upon one integration the "tardigrade slant" $y_1 = (A/a^2)[at - (1 - e^{-at})]$ with $\Lambda = A/[s(s+a)]$; upon a delay of interval b, the "delayed tardigrade" $y_1 = (A/a)[1 - e^{-a(t-b)}]x_1(t-b)$ with $\Lambda = Ae^{-bs}/(s+a)$.

The offset sinusoid and the offset damping, when differentiated, retain their characteristic oscillations, but are shifted to sym-

metry about the time axis, and changed in amplitude. Delay gives them a simple horizontal shift; integration produces horizontal shift, vertical shift, and change of amplitude.

Two Combining Methods

The two combining methods are as follows:

1 Superposition, giving the sum or difference of curve ordinates, or of the transferences.

2 Convolution, giving the product of transferences.

As a matter of fact, we have tacitly made use of superposition in setting up the two-term basic responses 2, 3, and 4 of Table 1, and the modifications already applied to these patterns might be thought of as convolution, since their transferences were multiplied by s, $1/s$, or e^{-bs} to form the new Λ. Much more complicated patterns, however, can be brought within our classification if we recognize superposition and convolution as general methods of combining the basic or modified types. A few of the possible curves which result are shown in Fig. 5. In naming them, the phrase "degree n" denotes the degree of the polynomial in s which forms the denominator of the transference, after clearing of fractions.

Superposition and convolution, used singly or conjointly, allow relatively simple procedures for predicting the indicial responses of more extended segments from the y_1 or Λ-values of the component members. We shall next show these methods for members manifolded, members sequenced, and members looped, and extend them to segments involving reflexed action and to arbitrary disturbances.

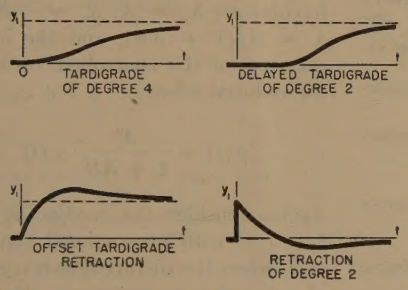


FIG. 5 SOME INDICIAL RESPONSES OBTAINED BY COMBINATION

Members Manifolded

The simplest case of combination (superposition used alone) arises when transfer members are manifolded and the disturbance has the offset form $x_1(t)$. Members are said to be "manifolded" (the term is taken from piping practice) when they are so connected as to receive identical and simultaneous disturbances, and to deliver their individual responses additively or subtractively to the segment below, both actions cascading. It should be noted that this is not always equivalent to the electrical "connection in parallel"; for the latter it is necessary that the identical variables be voltages and that the additive ones be currents, while cascading action is found less often than reflexing. An example of manifolding is the typical controller whose components give proportional-plus-floating action. Its indicial response is $y_1 = A + Bt$, and its transference is $\Lambda = A + B/s$, an offset and a slant superposed. Other examples appear in Table 2.

Members Sequenced

Convolution used alone is called for when transfer members are "sequenced," i.e., connected so that a cascaded disturbance from above develops in each a response which cascades to its neighbor below. If there are n such members, having transferences $\Lambda_1, \Lambda_2, \ldots \Lambda_n$, it can be shown that the transference for the combination is the product $\Lambda = \Lambda_1 \cdot \Lambda_2 \ldots \Lambda_n$, and that the indicial response at the last member, to the unit offset $x_1(t)$ acting on the first member, is $y_1 = \mathcal{L}^{-1}[(1/s)\Lambda]$. The Λ's may be positive or negative, and the order of sequencing is immaterial. It should be noted that sequencing is not always equivalent to the electrical "connection in series"; for the latter there must be an identical current in each member, and cascading is found less often than reflex action. As an example of convolution, consider the sequencing of two tardigrade members having transferences $\Lambda_a = a/(s + a)$ and $\Lambda_b = b/(s + b)$, with indicial responses $y_a = 1 - e^{-at}$ and $y_b = 1 - e^{-bt}$. By multiplying transferences we get for the sequence

$$\Lambda = \frac{a}{s + a} \cdot \frac{b}{s + b}$$

whence

$$y_1 = 1 + \frac{be^{-at}}{a - b} - \frac{ae^{-bt}}{a - b}$$

a tardigrade of degree 2. Sequenced members are very common, and further examples appear in Table 3.

TABLE 2 MEMBERS MANIFOLDED

	Transference	Response	Name
(1)	$B + Cs + D/s$	$B + Cx_1' + Dt$	Offset plus spike plus slant
(2)	$\frac{a}{s + a} + \frac{s}{s + a}$	$1 - e^{-at} + e^{-at} = 1$	Unit offset
(3)	$A - Ae^{-as}$	A for $0 < t < a$ 0 for $t > a$	Pulse
(4)	$\frac{s}{s^2 + \omega^2} + \frac{s\,e^{-\pi s/\omega}}{s^2 + \omega^2}$	$\omega^{-1}\sin \omega t$ when $t \leq \pi/\omega$ 0 when $t \geq \pi/\omega$	Interfering sinusoids
(5)	$\frac{a}{s + a} - \frac{b}{s + b}$	$e^{-bt} - e^{-at}$	Tardigrade retraction
(6)	$\frac{a}{s + a} + \frac{b}{s + b} - \frac{c}{s + c}$	$1 + e^{-ct} - e^{-bt} - e^{-at}$	Offset tardigrade retraction

Arbitrary Disturbances

One of the principal uses of the indicial response and its associated transference is to predict the response $y(t)$ of a transfer member to any arbitrary disturbing function $x(t)$, when $y_1(t)$ and Λ are known. It can be shown that the result is the derivative of a convolution

$$y(t) = \frac{d}{dt}[x * y_1] = \frac{d}{dt}\int_0^t x(t - \xi) \cdot y_1(\xi)d\xi$$

Sometimes this third member is readily evaluated just as it stands, but frequently it is easier to take transforms of the first equality, obtaining $\tilde{y} = s \cdot \tilde{x} \cdot \tilde{y}_1 = \tilde{x}\Lambda$, then $y(t) = \mathcal{L}^{-1}[\Lambda \cdot \mathcal{L}(x)]$. To illustrate, let a member with the indicial response $y_1 = 1 - e^{-bt}$ and transference $\Lambda = b/(s + b)$ be subjected to the cascaded disturbance $x = At$. Here $\mathcal{L}(x) = A/s^2$ and

$$y = \mathcal{L}^{-1}\left[\frac{A}{s^2} \cdot \frac{b}{s + b}\right] = \frac{A}{b}\left[bt - (1 - e^{-bt})\right]$$

That same response would be obtained if we imposed $x_1(t)$ on a sequence of two members having indicial responses $y_a = At$ and $y_b = 1 - e^{-bt}$. Thus, in the examples of sequencing in Table 3, one of the s-function factors may represent the transference of a member; the other, when divided by s and returned to the time domain, an arbitrary disturbance acting on that member.

TABLE 3 MEMBERS SEQUENCED

	Transference	Response	Name
(1)	$A \cdot B$	$AB\, x_1(t)$	Offset
(2)	$Ae^{-as} \cdot \frac{B}{s}$	$AB[t - a]\, x_1(t - a)$	Delayed slant
(3)	$As \cdot \frac{B}{s}$	$AB\, x_1(t)$	Offset
(4)	$\frac{A_1}{s} \cdot \frac{A_2}{s} \cdots \frac{A_n}{s}$	$\frac{A_1 \cdot A_2 \cdot \ldots A_n t^n}{n!}$	Parabola, degree n
(5)	$\frac{A}{s}[B + Cs + D/s$	$A[Bt + C + Dt^2/2$	Offset plus slant plus parabola
(6)	$A \cdot \frac{B}{s + a}$	$\frac{AB}{a}(1 - e^{-at})$	Tardigrade
(7)	$Ae^{-as} \cdot \frac{Be^{-bs}}{s + c}$	$\frac{AB}{c}[1 - e^{-c(t-a-b)}]x_1(t - a - b)$	Delayed tardigrade
(8)	$As \cdot \frac{B}{s + b}$	ABe^{-bt}	Retraction
(9)	$\frac{A}{s} \cdot \frac{B}{s + b}$	$\frac{AB}{b^2}[bt - (1 - e^{-bt})$	Tardigrade slant
(10)	$\frac{a}{s + a} \cdot \frac{a}{s + a}$	$1 - e^{-at}(1 + at)$	Tardigrade, degree 2
(11)	$\left(\frac{a}{s + a}\right)^n$	$1 - e^{-at}\left[1 + at + \ldots + \frac{(at)^{n-1}}{(n-1)!}\right]$	Tardigrade, degree n
(12)	$\frac{a_1}{s + a_1} \cdot \frac{a_2}{s + a_2} \cdots \frac{a_n}{s + a_n}$	$1 - \sum_{k=1}^{n} \lim_{s \to -a_k} \frac{(s + a_k)\, a_1 a_2 \cdots a_n e^{-akt}}{(s + a_1)(s + a_2) \cdots (s + a_n)}$	Tardigrade, degree n
(13)	$\frac{a}{s + a} \cdot \frac{b}{s + b} \cdot e^{-cs}$	$1 + \frac{b}{a - b} \cdot e^{-a(t-c)} - \frac{a}{a - b} \cdot e^{-b(t-c)}$ for $t > c$, 0 for $t < c$	Delayed tardigrade, degree 2; often a satisfactory approximation for tardigrade of degree n.
(14)	$\frac{As}{s + a} \cdot \frac{B}{s + a}$	$AB\, te^{-at}$	Tardigrade retraction
(15)	$\frac{As}{s + a} \cdot \frac{B}{s + b}$	$\frac{AB}{b - a}[e^{-at} - e^{-bt}$	Tardigrade retraction
(16)	$\left(\frac{s}{s + a}\right)^n$	$e^{-at} \cdot \sum_{k=0}^{n-1} \frac{(n-1)!(-at)^k}{(n-1-k)!(k!)^2}$	Retraction, degree n
(17)	$\frac{s}{s + a_1} \cdot \frac{s}{s + a_2} \cdots \frac{s}{s + a_n}$	$\sum_{k=1}^{n} \lim_{s \to -a_k} \frac{(s + a_k)s^{n-1} e^{st}}{(s + a_1)(s + a_2) \cdots (s + a_k)}$	Retraction, degree n
(18)	$Ae^{-as} \cdot \frac{Bs}{s^2 + \omega^2}$	$AB\omega^{-1} \sin \omega(t - a)$ when $t > a$, 0 when $t < a$	Delayed sinusoid
(19)	$Ae^{-as} \cdot \frac{B}{(s + \delta)^2 + \omega^2}$	$\frac{AB}{\delta^2 + \omega^2} - \frac{ABe^{-\delta(t-a)}}{\omega\sqrt{\delta^2 + \omega^2}} \cdot \cos[\omega(t - a) - \tan \delta/\omega]$ when $t > a$; 0 when $t < a$	Delayed offset damping
(20)	$As \cdot \frac{Bs}{(s + \delta)^2 + \omega^2}$	$ABe^{-\delta t}\sqrt{1 + (\delta/\omega)^2} \cdot \sin(\omega t + \tan^{-1} \omega/-\delta)$	Damping
(21)	$\frac{A}{s} \cdot \frac{B}{(s + \delta)^2 + \omega^2}$	$\frac{AB}{\delta^2 + \omega^2}\left[t - \frac{2\delta}{\delta^2 + \omega^2} + \frac{e^{-\delta t}}{\omega} \sin(\omega t - 2 \tan^{-1} \omega/-\delta)\right]$	Slant damping
(22)	$\frac{As}{s^2 + \omega^2} \cdot \frac{Bs}{s^2 + \xi^2}$	$\frac{AB}{\xi^2 - \omega^2}[\cos \omega t - \cos \xi t]$	Sinusoid, degree 4
(23)	$\frac{As}{s^2 + \omega^2} \cdot \frac{B}{s^2 + \omega^2}$	$\frac{AB}{2\omega^3}(\sin \omega t - \omega t \cos \omega t)$	Unstable sinusoid

Members Looped

Influence-path segments are often connected in a closed cause-effect circuit called a control loop, for the purpose of obtaining either substantial constancy of a responding variable or its prompt conformity to a command signal. We distinguish two parts of the loop: a "forward" segment, extending from the point where a disturbance acts on the loop to the point where the principal response is taken off; and a "feedback" segment, which receives the latter signal and converts it into a secondary response which is fed into the forward segment along with the external disturbance.

Quite frequently both segments are bounded by cascading actions, and in this case a simple formula relates the various signals. Let $x_1(t)$ be the external disturbance (whether an intrusion to be annulled or an order to be followed), Λ_y the transference of the forward segment, $v_1(t)$ the closed-loop or principal response, Λ_z the transference of the feedback, and $z_1(t)$ the feedback response arising from v_1. Now $\tilde{z}_1 = \tilde{v}_1 \cdot \Lambda_z$ and $\tilde{v}_1 = (\tilde{x}_1 + \tilde{z}_1)\Lambda_y$, where $\tilde{x}_1$ and $\tilde{z}_1$ must have opposing signs if the loop is to provide corrective action. Then $\tilde{v}_1 = \tilde{x}_1\Lambda$, where $\Lambda = \Lambda_y/(1 - \Lambda_y\Lambda_z)$. Convolution (or another superposition) may be used to form Λ_y and Λ_z if forward and feedback segments consist of sequenced (or manifolded) members. It is sometimes convenient to consider the product $\Lambda_y\Lambda_z$ as a transference for the several cascading segments connected in sequence. For this purpose the loop can be opened between any two of these, an experimental disturbance being applied above one and its response measured below the other.

As an example of looping, consider a plant defined by the response $Ax_1(t)$ operating with a controller defined by $-Bv_1(t)$, when disturbance $x_1(t)$ acts on the plant. The respective transferences are $\Lambda_y = A$, $\Lambda_z = -B$, and $\Lambda = A/(1 + AB)$, and the indicial response of the controlled variable is the reduced offset

$$v_1(t) = \frac{A}{1 + AB} \cdot x_1(t)$$

Again, consider the tardigrade plant $1 - e^{-at}$ with proportional controller $-B$, when the disturbance is a change of set point or program. Here the forward segment includes both controller and plant, with $\Lambda_y = (-aB)/(s + a)$, and the indicial response of the feedback segment degenerates to $z_1 = 1$, so that

$$\Lambda_z = 1, \Lambda = \frac{-aB}{s + a + aB}, \text{ and}$$

$$v_1 = \frac{-B}{1 + B}[1 - e^{-a(1+B)t}]$$

Further examples are given in Table 4.

Reflex Action

The response of a transfer segment is said to be "reflexed" if its behavior depends simultaneously upon the disturbance above, which originated it, and on the response of the segment below, which it in turn disturbs. Frequently the two effects oppose one another, providing an inherent sort of control which is called "self-regulation." Obviously, the behavior of variables linked by reflex action is indeterminate until fixed by known behaviors at points of cascading action which one reaches eventually by proceeding in all possible directions along the network. However, reflex action with segments of finite size produces no forms of indicial response which cannot be obtained by modifying or combining the four basic patterns as described.

The simultaneous differential equations which state the physical relations between variables and constants within the larger cascade-bounded segment, the member, will frequently be found

TABLE 4 MEMBERS LOOPED

	Λ_y	Λ_z	Λ	Response	Name
(1)	$-A$	B	$\frac{-A}{1+AB}$	$\frac{-A}{1+AB}$	Offset
(2)	As	$-B$	$\frac{As}{1+ABs}$	$\frac{1}{B}\cdot e^{-t/AB}$	Retraction
(3)	A	$-Bs$	$\frac{A}{1+ABs}$	$A(1-e^{t/AB})$	Tardigrade
(4)	As^2	$-B$	$\frac{As^2}{1+ABs^2}$	$\frac{1}{B}\cos \omega t$ with $\omega^2 = 1/AB$	Sinusoid
(5)	A	$-Bs^2$	$\frac{A}{1+ABs^2}$	$A(1-\cos \omega t)$ with $\omega^2 = 1/AB$	Offset sinusoid
(6)	A	$-B/s$	$\frac{A}{1+AB/s}$	Ae^{-ABt}	Retraction
(7)	A/s	$-B$	$\frac{A/s}{1+AB/s}$	$\frac{1}{B}(1-e^{-ABt})$	Tardigrade
(8)	$\frac{A}{s+a}$	$-B$	$\frac{A}{s+a+AB}$	$\frac{A}{b}(1-e^{-bt})$ with $b = a + AB$; if $B >> a/A$ controlled response is faster and balances out at a smaller offset	Tardigrade
(9)	$\frac{A}{s^2+\omega^2}$	$-B$	$\frac{A}{s^2+\omega^2+AB}$	$\frac{A}{\omega_1^2}(1-\cos \omega_1 t)$ with $\omega_1^2 = \omega^2 + AB$; if $B >> \omega^2/A$, controlled response is a more rapid oscillation, of smaller amplitude	Offset sinusoid
(10)	Ae^{-as}	$-B$	$\frac{A}{e^{as}+AB}$	$A[x_1(t-a) - bx_1(t-2a) + \ldots + (-b)^{n-1}x_1(t-na)]$	

where $b = AB$, $n = 1, 2, \ldots \infty$; if $AB < 1$ $v_1(t)$ is an offset damping which ultimately approaches $A/(1+AB)$ through a sequence of steps whose ordinates oscillate about $A/(1+AB)$.

(11)	$\frac{A}{(s+a)^2+b}$ where $\omega^2 = b + AB$	$-B$	$\frac{A}{(s+a)^2+b+AB}$	$A\left[\frac{1}{a^2+\omega^2} - \frac{e^{-at}}{\omega\sqrt{a^2+\omega^2}}\cos(\omega t - \tan^{-1} a/\omega)\right]$	Offset damping
(12)	$\frac{A}{s+a}$ where $\delta = a/2$, $\omega^2 = AB - \delta^2 \neq 0$	$\frac{-B}{s}$	$\frac{As}{s^2+as+AB}$	$A/\omega\, e^{-\delta t}\sin \omega t$	Damping
(13)	$(A)\ (-B)$	1	$\frac{-AB}{1+AB}$	$\frac{-AB}{1+AB}$	Offset
(14)	$(A/s)\ (-B)$	1	$\frac{-AB}{s+AB}$	$-[1-e^{-ABt}]$	Tardigrade
(15)	$\frac{\left(\frac{a}{s+a}\right)\left(\frac{b}{s+b}\right)(-c)}{1+\frac{abc}{(s+a)(s+b)}}$ where $2\delta = a + b$, $\omega^2 = ab(1+2c) - (a+b)^2/4$	1	$\frac{-abc}{(s+\delta)^2+\omega^2}$	$-\frac{abc}{\delta^2+\omega^2} - \frac{abce^{-\delta t}}{\omega\sqrt{\delta^2+\omega^2}}\cos(\omega t - \tan^{-1}\delta/\omega)$	Offset damping

more convenient than the corresponding transferences as a starting point for computing the behavior of these variables. Such equations, based either upon loops or upon nodes, may be solved by classical methods, but the solution becomes simple algebra through use of $\mathcal{L}$-transforms. As an illustration, consider the electrical network known as an RC ladder, having two sections

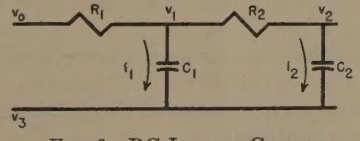


FIG. 6 RC LADDER CIRCUIT

as shown in Fig. 6, with cascading action above $v_0 - v_3$ and below $v_2 - v_3$, $v_0 - v_3 = 1$ for $t > 0$, and all variables (and their derivatives and integrals) zero for $t < 0$. Voltage differences are

$$v_0 - v_1 = i_1 R_1,\ v_1 - v_3 = \frac{1}{C_1}\int_0^t (i_1 - i_2)dt$$

$$v_1 - v_2 = i_2 R_2,\ v_2 - v_3 = \frac{1}{C_2}\int_0^t i_2 dt$$

which become upon taking $\mathcal{L}$-transforms and combining

$$\begin{cases} 1/s = \tilde{i}_1 R_1 + i_1/C_1 s - \tilde{i}_2/C_1 s \\ 0 = \tilde{i}_2 R_2 + \tilde{i}_2/C_2 s - \tilde{i}_1/C_1 s + \tilde{i}_2/C_1 s \end{cases}$$

These algebraic equations can readily be solved for $\tilde{i}_1$ and $\tilde{i}_2$, and from the $\mathcal{L}^{-1}$ transforms the currents, voltages, and voltage drops are obtainable.

It is possible, however, to compute an internally reflexed member by combined convolution and superposition. Suppose one of its segments is acted on by a disturbance w_f, and yields a reflexed response w_g which acts on an adjoining segment having the response w_h. Then the definition of reflex action gives us the $\mathcal{L}$-transform relation

$$\tilde{w}_g = \tilde{w}_f[\Lambda(\tilde{w}_g/\tilde{w}_f] + \tilde{w}_h[\Lambda(\tilde{w}_g/\tilde{w}_h]$$

There is a self-regulation of w_g if it turns out that the two terms here have opposing signs. We proceed along the network, writing similar expressions for reflexed variables above and below w_g, and eventually encounter cascading actions. Upon taking $\mathcal{L}$-transforms of the entire set, we obtain enough simultaneous algebraic equations to give a solution for any variable desired. To illustrate, consider the fluid storage member of Fig. 7, with capacitance C, fed through piping of resistance R and discharged by a positive-displacement pump; suppose that $\theta_0 = 1$ and $q_2 = 0$ for $t > 0$, and that all variables (and their derivatives and integrals) are zero for $t < 0$. In this situation the transferences in

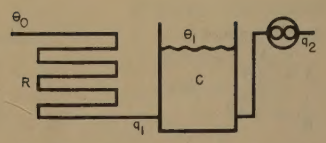


FIG. 7 FLUID STORAGE SYSTEM

action will be $\Lambda(\tilde{q}_1:\tilde{\theta}_0) = 1/R$, $\Lambda(\tilde{q}_1:\tilde{\theta}_1) = -1/R$, $\Lambda(\tilde{\theta}_1:\tilde{q}_1) = 1/Cs$. Upon using the $\tilde{w}_o$ relation twice, we obtain $\tilde{q}_1 = \tilde{\theta}_0/R - \tilde{\theta}_1/R$ and $\tilde{\theta}_1 = \tilde{q}_1/Cs$, from which $\theta_1 = 1 - e^{-t/RC}$ and $q_1 = (1/R)e^{-t/RC}$.

CONCLUSION

Alternative schemes of classification deserve some brief comment. One might be based upon the form of the ordinary differential equation, but this neglects behaviors described by partial differential, integro-differential, and difference-differential equations, and is ambiguous because more than one equation form leads to the same behavior pattern. Another might use the frequency spectra of amplitude ratio and lag angle, based upon sinusoidal disturbances at many frequencies, but the author believes these are more difficult for the engineering student and well-nigh unintelligible to the mechanic. A third might be built around the transfer locus, but this is not readily grasped until the student has considerable practice in manipulating vectors. Still another might use the $\mathcal{L}$-transform without multiplication by s, but this is less convenient in working with connected segments.

The underlying concept of the scheme proposed herein is that of a time behavior resulting from a simple disturbance and coded by a quantitative expression in s. Four basic curve patterns, together with three methods of modifying them and two methods of combining them, are deemed adequate for the linear dynamic behaviors met in the engineering of automatic controls. Considerable thought was given to the terms used, and to their precise definition, but the specific names and letter symbols used are not essential to the scheme, and perhaps better ones will be brought out by discussion. Graphic forms are advantageous because they can be widely understood, readily obtained from experiment or theory, and descriptively named. By associating with each the expression $s\cdot\mathcal{L}(y_1)$, provision is made for exactness of specification, convenience of manipulation, and order of arrangement. The examples given illustrate the application to single members, to groups sequenced or manifolded or looped, to members internally reflexed, and to arbitrary disturbances.

ACKNOWLEDGMENT

Acknowledgment is made of helpful comments made concerning the manuscript by Geraldine A. Coon and by George A. Philbrick. The work was undertaken at the instance of the IIRD Committee on Theory.

Discussion

GERALDINE COON.[3] The indicial response is a logical, but not always practical, basis for classifying the linear transfer members of ordinary occurrence in automatic control. Inasmuch as the transference often involves a number of time constants, the detailed computation of the indicial response may be a laborious matter. In the case of a system with feedback, the factoring of $1 - \Lambda_y\Lambda_z$ may be a time-consuming, but necessary, step in obtaining such a response. However, substantial information concerning the stability of the system may be obtained readily from the transfer locus or *db*-log ω plots without such detailed calculations.

The names "offset," "delayed," and "sinusoid" are quite acceptable; "tardigrade" and "retraction" are descriptive, but more suitable terms should be sought. There is no apparent justification for renaming the "impulse function" and "ramp" which are already in use. Some of the combinations of names are open to question. For example, response (17) of Table 3, called a tardigrade retraction of degree n, seems to be a retraction of degree n; response (9) of Table 3, the tardigrade slant, is actually the difference rather than the sum of a slant and a tardigrade as might be expected.

Some criticism may be leveled at the definitions of equilibrium and transference. The initial conditions should not be included in the definition of equilibrium; rather it is preferable to associate them with the transference of the system. Thus, transference may be defined as the ratio of the Laplace transforms of any response and the input that produces it under the condition that the system starts from rest when the input is first applied.

AUTHOR'S CLOSURE

Looped systems with $\Lambda = \Lambda_y/(1 - \Lambda_y\Lambda_z)$ can be shown to lie within the stability range by use of the Routh or Hurwitz criteria based upon the constant coefficients of $1 - \Lambda_y\Lambda_z$, without the necessity of factoring. When this function is a polynominal of degree higher than third, it is not evident, without factoring, just how far below sustained oscillation the nth degree indicial response will lie. If close, it will be a damping; if distant, a tardigrade. In most practical cases of looped systems, it probably will be a damping, and might be so designated when neither the factors nor the experimental curves are at hand. It is true that the transfer locus, obtainable after substituting $j\omega$ for s in $\Lambda_y\Lambda_z$, is somewhat more informative on this point, and this is a valid argument for its use. The indicial response still seems preferable, however, to introduce the student to types of transfer members.

It was hoped that better terms might be suggested for those unfavorably received, but no substitutions have been offered. Response (17) of Table 3 is indeed a retraction of degree n, and has been corrected. The argument for naming (9) of Table 3 is that it shows a tardigrade approach to what eventually becomes a constant slope. "Spike" was proposed as more descriptive of the pattern than is "impulse." When there are horizontals at both ends of a slope, as in Fig. 2 of the paper, the combination seems more properly a "ramp" than when the slope continues indefinitely, as in $y_1 = At$.

Miss Coon's definition of transference is acceptable, except that the "start from rest" is too restrictive. In speed regulation, for example, deviations may be measured from a state of equilibrium defined by uniform rotation.

[3] Engineering Research Division, Taylor Instrument Companies, Rochester, N. Y.

NOTE (refer to footnote 2 of paper, page 408): The complex number Λ, as a function of s, is a generalized form of similar operational expressions in p or frequency relations in $j\omega$ used by many authors, e.g., G. S. Brown and A. C. Hall, "Dynamic Behavior and Design of Servomechanisms," Trans. ASME, vol. 68, 1946, pp. 503–524, or G. A. Philbrick, "Unified Symbolism for Regulatory Controls," Trans. ASME, vol. 69, 1947, pp. 47–67. Its definition in terms of $\mathcal{L}$- transforms gives it the advantages of greater rigor and wider effectiveness, as pointed out by M. F. Gardner and J. L. Barnes, "Transients in Linear Systems," John Wiley & Sons, Inc., New York, N. Y., vol. 1, 1942.